Mário A. Perini

PARA UMA NOVA GRAMÁTICA DO PORTUGUÊS

Sumário

Prefácio

Este livro surgiu do reconhecimento da necessidade de se elaborar uma nova gramática do português. Enquanto a "crítica da gramática tradicional" vai pouco a pouco conquistando um lugar nos cursos de Letras, sente-se agudamente a falta de alternativas viáveis: se a gramática tradicional é inadequada, o que colocar em seu lugar? O lingüista, ao criticar a gramática tradicional, freqüente mente se choca com a objeção irrespondível: é indispensável continuar trabalhando com uma gramática cujas deficiências são evidentes, pois não existe outra que se possa utilizar.

Não que não haja descontentamento em todas as áreas interessadas: os professores sentem que a doutrina gramatical é ultrapassada, incoerente e muitas vezes simplista até a ingenuidade; os alunos tendem a desencantar-se de uma disciplina que só tem a oferecer-lhes um conjunto de afirmações aparentemente gratuitas e sem grande relação com fatos observáveis. Na sala de aula, às vezes o único refúgio são as atitudes autoritárias, quando ó professor não se sente em condições de liderar discussões verdadeiramente racionais sobre gramática.

Ora, não se pode esperar do professor de Português, por mais bem formado que sej a, que vá elaborando sua gramática à medida que se desenvolve o curso; ele é um profissional do ensino, e não um lingüista: não tem o treinamento (nem o tempo) para levar a efeito pesquisa tão longa e complexa. A responsabilidade deve cair, sem dúvida, sobre os pesquisadores da universidade, das áreas de Lingüística e de Língua Portuguesa: a esses cabe a obrigação de criar a nova gramática.

As falhas da gramática tradicional são, em geral, resumidas em três grandes pontos: sua inconsistência teórica e falta de coerência interna; seu caráter predominantemente normativo; e o enfoque centrado em uma variedade da língua, o dialeto padrão (escrito), com exclusão de todas as outras variantes. Todos os três pontos merecem atenção cuidadosa; só teremos uma gramática satisfatória como base para o ensino quando os três estiverem devidamente repensados. Assim, a gramática deverá, primeiro, colocar em seu devido lugar as afirmações de cunho normativo: não necessariamente suprimindo-as, mas apresentando o dialeto como uma das possíveis variedades da língua, adequada em certas circunstâncias e inadequada em outras (é tão "incorreto" escrever um tratado de Filosofia no dialeto coloquial quanto namorar utilizando o dialeto padrão). Depois, a gramática deverá descrever pelo menos as principais variantes (regionais, sociais e situacionais) do português brasileiro, abandonando a ficção, cara a alguns, de que o português do Brasil é uma entidade simples e homogênea. Finalmente, e acima de tudo, a gramática deverá ser sistemática, teoricamente consistente e livre de contradições.

Ë este último aspecto do planejamento da nova gramática que nos ocupará aqui. Como é inevitável em um livro tão pequeno, fui obrigado a

selecionar alguns assuntos, deixando outros de lado; portanto, tomei como tópicos os que me parecem ser as grandes falhas da doutrina gramatical vigente. Em cada caso, parto de uma crítica da posição tradicional; em seguida, proponho as linhas gerais de uma solução. O resultado, espero, contribuirá para dar idéia de uma proposta para uma nova gramática do português.

Para isso, foi necessário abordar certo número de questões teóricas. Com efeito, é minha opinião que, se a gramática tradicional está em tão mau estado hoje em dia, isso se deve em grande parte a uma injustificada timidez em abordar os problemas gramaticais dentro de uma perspectiva teórica, ou seja, com a preocupação de generalizar. Alguns estudiosos de gramática chegam a exprimir certa desconfiança da teoria, que consideram desvinculada das questões reais de análise gramatical. Neste livro, parto da crença de que essa atitude é equivocada, e que não há, simplesmente, esperanças de se chegar a uma prática gramatical realmente racional — e, portanto, educacionalmente valiosa — sem uma fundamentação teórica suficiente.

O presente trabalho é parte de um projeto maior, que espero levar a efeito nos próximos anos, de elaboração de uma nova descrição do português padrão (uma nova gramática portuguesa). Mas a publicação da nova gramática, em si, cairá no vazio se não se fizer acompanhar de um amplo debate sobre os fundamentos do ensino gramatical: seus objetivos e as maneiras de atingi-los. Este livro deve, pois, ser encarado como um convite à discussão dos grandes traços de uma nova gramática portuguesa, que possa servir de apoio à renovação do ensino gramatical entre nós. Mais do que a substituição de uma doutrina gramatical por outra (o que seria de utilidade questionável), creio que se deve almejar a criação de novas atitudes, caracterizadas por maior responsabilidade teórica, maior rigor de raciocínio, libertação do argumento da autoridade — em uma palavra, mais espírito crítico. Só assim poderá o ensino da gramática proporcionar um campo para o exercício da argumentação e do raciocínio, contribuindo para a formação intelectual dos estudantes.

A leitura deste livro não pressupõe treinamento específico em Lingüística. Os termos técnicos utilizados são definidos no próprio texto, ou no vocabulário crítico final. Exige-se apenas conhecimento da gramática tradicional e uma mente aberta para a discussão de seus fundamentos. Isso não significa que o texto seja sempre fácil de ler. Há complexidades inevitáveis, em especial nos pontos em que a perspectiva adotada difere mais profundamente das idéias tradicionais. Mas certamente ninguém ignora a grande complexidade da linguagem humana, nem o estado de ignorância em que nos encontramos a esse respeito. Aqui, como em toda a parte, fugir ao problema não é maneira de começar a resolvê-lo. A tarefa que nos espera é longa e difícil; mais uma razão para que não a adiemos.

Alguns amigos me ajudaram na realização deste trabalho, de diversas formas. Em especial, mencionarei Mary Kato, Jânia R do Nascimento. A eles, e aos demais, meu muito obrigado.

1 Introdução

Todos concordam que é necessário descrever a língua em novos moldes. Antes, porém, de iniciar o trabalho, será preciso chegar a um acordo sobre que moldes serão esses. Lembremo-nos de que se trata de preparar uma gramática pedagógica: isso significa que uma das tarefas a enfrentar é a de selecionar (ou, mais provavelmente, inventar) uma linguagem para transmitir os resultados da investigação lingüística das últimas décadas sem, por um lado, falsificá-los, nem, por oqtro, tornar o texto inacessível a quem não seja um lingüista profissional.

Ë evidente que alguma "falsificação" será inevitável, na forma de simplificações, na forma de soluções inseguras, apresentadas ao lado de soluções bem fundamentadas, na forma sobretudo de uma certa escassez (nunca ausência completa!) de argumentação polêmica. Creio que essa situação é não apenas suportável, mas indispensável. Certa mente é assim que procedemos autores de textos de Física, Biologia ou Sociologia: ninguém tenta colocar já no primeiro momento, ou em textos dirigidos a não-especialistas, todas as complexidades da teoria, ou toda a confusão das discussões acadêmicas. O objetivo é, antes de tudo, comunicar resultados, deixar entrever métodos e evitar dar a impressão de que a disciplina é um conjunto de princípios fixos e universalmente aceitos.

Dentro das linhas acima indicadas, o primeiro problema é a escolha do arcabouço teórico a ser adotado. Ë desejável neutralizar, tanto quanto possível, os grandes problemas não-resolvidos da teoria lingüística moderna; apresentar-se-á, assim, uma visão propositalmente superficial, que permita a tomada de um "panorama geral" aceitável para muitos lingüistas. Como um exemplo, tome mos o problema fundamental da relação entre estrutura observável (superficial) e interpretação semântica (o significado). Há no momento algumas dezenas de soluções concorrentes, todas insuficientes, para esse problema; optar por uma delas, arbitrariamente, seria pretender que todos aprendessem os detalhes de um modelo (com todo o trabalho que isso implica) sem nenhuma garantia de que seja, ou venha a ser, um modelo predominante em Lingüística (isso se deve exigir, é claro, de lingüistas profissionais; mas nunca de professores de línguas, cujo interesse em Lingüística -embora grande, é instrumental).

Creio que existe a possibilidade de uma solução de compromisso segundo as linhas seguintes: incluiremos na gramática uma dupla descrição, a saber (a) uma descrição em termos formais da estrutura sintática superficial; e (b) uma descrição de aspectos da interpretação semântica, colocada, na medida do possível, em paralelo com a descrição sintática. No caso da sintaxe e da morfologia, far-se-á um esforço no sentido de preservar, onde possível, a nomenclatura tradicional. Já no caso da semântica não vejo como isso possa ser feito, dada a inexistência pura e simples de uma terminologia semântica minimamente coerente dentro da gramática tradicional.

O grau de exaustividade desses dois componentes será necessariamente muito diferente, em vista do atual estado dos estudos

lingüísticos. Parece-me indicado fazer da descrição formal (morfossintática) a linha mestra da descrição, e apenas nesse aspecto se tentará dar uma visão razoavelmente abrangente da estrutura da língua. Quanto ao componente semântico, será preciso estabelecer metas mais modestas: ele será esboçado, procurando-se generalizações, mas sem oferecer propriamente um conjunto estruturado. A semântica terá de ser, a rigor, uma espécie de antologia de apêndices à descrição gramatical. Não vejo inconveniente nessa solução, que me parece a melhor nas atuais circunstâncias.

Já que convém, dados os objetivos pedagógicos da gramática, conservar tanto quanto possível a linguagem tradicional, faz sentido iniciar a discussão explicitando os pontos fracos da doutrina gramatical vigente. Procurarei isolar neste trabalho as grandes questões básicas que não são adequadamente tratadas dentro da gramática tradicional (de agora em diante, abreviadamente GT). Minha abordagem é teórica: preocupo-me em explicitar as crenças e princípios gerais que subjazem à prática da análise. Nisso o presente trabalho contrasta com a maioria das obras correntes; com efeito, é incomum discutir-se as bases teóricas da GT: a maior parte dos trabalhos se concentram na análise de problemas específicos (uma rara exceção é Hauy, 1983). O resultado, como tentarei mostrar, é uma gramática construída sobre um caos teórico, não sendo de admirar que ela seja também caótica.

Antes, gostaria de fazer uma advertência: nos exemplos aqui citados não é minha intenção acusar os autores das gramáticas atuais de incompetência ou de falta de honestidade intelectual. Eles são, e creio que muitos deles o reconhecem, vítimas de uma tradição que fez da tarefa do gramático pouco mais do que uma compilação estéril. Acontece que, até bem recentemente, havia pouca alter nativa fora dessa linha; faltavam trabalhos de análise da língua feitos sobre bases sólidas, e os próprios princípios da Lingüística moderna eram, por assim dizer, privilégio de uns poucos eleitos. A partir do grande desenvolvimento dos estudos lingüísticos no Brasil que presenciamos nos últimos anos tornou-se possível pensar realisticamente na elaboração de uma alternativa à gramática usual. A crítica aqui feita, portanto, é uma crítica à gramática, não aos gramáticos.

Um exemplo

Voltemos agora ao nosso tema principal. Uma definição comum de "sujeito" é a seguinte:

(1) "O sujeito é o termo sobre o qual se faz uma declaração". (CUNHA, 1975, p. 137.)

A partir dessa definição podemos tirar um exemplo da pouca consistência da GT. (1) é a única definição de sujeito dada na gramática; é de se esperar, pois, que ela reflita a noção de sujeito válida para toda a análise. Quero dizer: no momento em que apresentamos (1) como a definição de sujeito, assumimos o compromisso de mantê-la como a definição de sujeito em toda a gramática. Em outras palavras, o termo "sujeito" corresponde a uma noção unificada e consistente, à qual as regras gramaticais podem fazer referência. De outra maneira, para começar, não se compreenderia a necessidade de definir sujeito.

Ora, logo adiante na mesma gramática, encontramos a afirmação seguinte:

(2) “Algumas vezes o verbo não se refere a uma pessoa determinada, ou por se desconhecer quem executa a ação, ou por não haver interesse no seu conhecimento. Dizemos, então, que o sujeito é indeterminado”. (CUNHA, 1975, p. 141.)

Já aqui se desrespeita a definição dada algumas páginas antes. Se é que existe sujeito indeterminado, ele deveria ocorrer nos casos em que se desconhece o ser sobre o qual se faz a declaração; afinal, a definição de sujeito só menciona essa característica dos sujeitos. No entanto, (2) está formulada como se o sujeito tivesse sido definido em termos de quem pratica a ação.

A gramática, nesse ponto, é inconsistente, ou pelo menos incompleta (não menciona a presumível conexão entre o sujeito e o ser que pratica a ação). E esse não é um caso isolado; é um exemplo dentre muitos, decorrências de uma atitude muito generalizada de falta de um esforço teórico constante. Ao se enunciar uma afirmação gramatical como (2), é preciso estar consciente de certas crenças que subjazem à afirmação; no caso de (2), está sub jacente a crença de que o sujeito exprime o nome do ser que pratica a ação. O esforço teórico a que me refiro consiste em tentar conciliar essas crenças subjacentes em um corpo doutrinário logicamente consistente — uma teoria gramatical. E isso falta na GT.

Se tivesse de apontar a grande falha fundamental da nossa tradição gramatical, eu escolheria justamente essa: a ausência de conscientização adequada do importe teórico das afirmações que constituem a gramática. Esse problema está na raiz do divórcio entre a doutrina gramatical, tal como explicitada em definições como (1) ou (2), e a prática da análise, tal como se observa quando um exemplo concreto é considerado (ver a seção 2). E está na raiz também da baixa qualidade de nossas descrições gramaticais. No momento em que assumimos a responsabilidade de vincular nossa análise, e toda ela, a uma teoria geral do funcionamento da língua, os defeitos ficam evidentes para qualquer um. Como disse acima, somos todos vítimas de uma tradição deformada; e sua maior deformação é precisamente o tipo de irresponsabilidade teórica que acabamos de ver.

A tarefa de fazer a nova gramática principia, por conseguinte, por uma conscientização das deformações da doutrina e da prática gramaticais. Nesse trabalho não pre- Doutrina explícita tendo restringir-me à crítica, tão freqüente e sempre fácil demais, da GT; mas vou começar pela crítica, tendo a preocupação de isolar os grandes problemas gerais. Espero que a partir daí seja possível começar a colocar novos alicerces, com o objetivo de construir algo de novo no lugar do que for eventualmente descartado.

2 doutrina explicita e doutrina implícita

Voltemos à definição de sujeito, a saber:

(1) “O sujeito é o termo sobre o qual se faz uma declaração”.

Já vimos que, em outras passagens da gramática, uma outra definição de sujeito (o termo que denota o ser que pratica a ação) aparece debaixo do pano. Agora vamos examinar alguns exemplos concretos de análise. Digamos que se peça a uma pessoa gramaticalmente treinada para identificar os sujeitos das orações abaixo:

(3) Carlinhos corre como um louco

(4) Carlinhos machucou Camilo

(5) esse bolo eu não vou comer

(6) em Belo Horizonte chove um bocado

Ela dirá que o sujeito de (3) e de (4) é Carlinhos; o de (5) é eu; e (6) não tem sujeito. Essas análises estão de acordo com a prática corrente, e creio que são de aceitação universal. Mas até que ponto se harmonizam com a definição (1), também geralmente aceita? Ë surpreendente verificar como são numerosos os choques entre a definição e a análise. Em (3) pode-se dizer sem problemas que a oração veicula uma declaração sobre Carlinhos, e sobre ninguém mais. Já em (4) isso não fica assim tão evidente:

não haverá aí também uma declaração sobre Camilo? Quando chegamos a (5) a situação se torna ainda mais desconfortável: como defender a tese de que (5) é uma afirmação acerca de mim, e não acerca do bolo? Final mente, (6), que é uma oração sem sujeito, necessariamente (segundo a definição) não deveria estar declarando nada sobre coisa alguma; no entanto, (6) exprime claramente uma declaração sobre Belo Horizonte.

Isso exemplifica uma contradição bastante clara entre a definição explícita de sujeito e a prática de identificação de sujeitos tal como se encontra quando da análise de casos concretos. Tais contradições são, em geral, toleradas, e mesmo ignoradas, por aqueles que trabalham com a GT. A razão para isso é, a meu ver, a seguinte: existe na verdade uma dualidade de doutrinas gramaticais dentro do que chamamos gramática tradicional. Uma dessas doutrinas está expressa, mais ou menos, nas gramáticas usuais. Essa doutrina (a que podemos chamar "doutrina gramatical explícita", ou-DGEx é que inclui definições como (1), acima, que conceitua o sujeito como o termo sobre o qual se faz uma declaração. Sabemos, porém, que as mesmas pessoas que propõem ou aceitam tais definições não as seguem na prática. Quando enfrentam a tarefa de encontrar o sujeito de uma oração, sempre analisam orações com chover como não tendo sujeito, muito embora, como vimos, tais orações possam perfeitamente fazer declarações sobre alguma coisa.

Ê que existe aqui, subjacente a essa análise, uma outra definição de sujeito, que aplicamos quando quere mos encontrar o sujeito de uma oração. Ë importante frisar que essa definição implícita é tão bem conhecida, embora em nível não-consciente, quanto (1). Se tomarmos algumas pessoas com instrução gramatical, elas terão tanta facilidade em citar (1) quanto em encontrar o sujeito de (5) ou (6). Essas pessoas dirão que o sujeito de (5) é eu, e dirão que (6) não tem sujeito, em geral sem perceber que isso não pode ser feito com base em (1).

Tudo se passa como se a aprendizagem da gramática envolvesse duas tarefas não-relacionadas. Primeiro, é preciso aprender, entre outras coisas, a identificar o sujeito de uma oração; depois, é preciso aprender a definição de sujeito. Mas note-se: a identificação dos sujeitos não se faz com base na definição aprendida; faz-se com base em alguma outra definição, nunca exteriorizada, mas inegavel mente existente (por que senão, como explicar o alto nível de coincidência entre as diversas pessoas ao identificarem os sujeitos das orações?). Aprendemos (1), mas não a levamos a sério. Identificamos o sujeito da maneira que sentimos ser a mais adequada, muito embora isso nos coloque em choque com a DGEx., de que (1) é uma parte. Ou seja, aprendemos a identificar o sujeito apesar do nosso conhecimento da DGEx. (que aqui nos aponta um caminho que não seguimos), e não por causa dele. Para explicar esse fato, postulo a existência de uma doutrina gramatical implícita (DGImp.), que não é nunca explicitada, nem reconhecida como existente, mas que na verdade guia nossas decisões dentro da prática da análise gramatical.[1]

De acordo com essa doutrina implícita, a definição de sujeito não pode ser (1). Qual será ela, então? No momento, não disponho de uma definição realmente completa e adequada a todos os casos; mas acho que a seguinte é uma aproximação:

(7) Sujeito é o termo com o qual o verbo concorda.

A prática gramatical observada sugere que (7) seja, pelo menos, parte da definição de sujeito existente na DGImp. Com efeito, (7) nos permite não só identificar Carlinhos como sujeito de (3) e de (4), mas também, inequivocamente, eu como sujeito de (5). Já (6) não tem no sujeito porque o verbo não concorda com nenhum dos termos da oração.

Em resumo: a prática dos estudos gramaticais revela a existência de duas doutrinas, a DGEx., que é “oficial mente reconhecida” e explicitada nas gramáticas, e a DGImp., nunca explicitada, mas que realmente subjaz à análise realizada. Essa não é uma descoberta nova. Creio que Jespersen se refere a esse fenômeno quando afirma:

o gramático treinado sabe se uma palavra dada é um adjetivo ou um verbo não por se referir a tais definições, mas praticamente da mesma maneira pela qual todos nós ao vermos um animal sabemos se é uma vaca ou um gato

(JESPERSEN, 1924, p. 62.)

Incoerência e autoritarismo

A DGEx. e a DGImp. só coincidem em parte; há grandes áreas em que elas entram em choque. No entanto, as análises levadas a efeito (com base na DGImp.) são usualmente j-ustif+eadas com base na DGEx. O resultado, como não podia deixar de ser, é um discurso incoerente, que muitas vezes

força o professor a assumir posições autoritárias ou de "dono da verdade" — por exemplo, negando a evidência, ou então dando a entender ao aluno que ele, professor, é detentor de conhecimentos mais avançados, que permitem a solução das incoerências presentes, mas que são por demais esotéricos para que o aluno a eles possa ter acesso.

Um exemplo de negação da evidência temos quando um aluno aponta (corretamente) que brilham, em:

(8) os diamantes brilham mais que os topázios exprime uma qualidade, e não uma ação, estado ou fenômeno; logo, não deveria ser considerado um "verbo", se é que verbo é "a palavra que exprime ação, estado ou fenômeno". O professor poderá responder que não, que em (8) os diamantes "praticam a ação de brilhar". Ora, tais respostas acabarão nos levando a uma definição do tipo "ação é tudo aquilo que se exprime por meio de um verbo" — o que esvazia a definição de verbo como palavra que exprime ação, pois a reduz a uma tautologia.

A atitude de "dono da verdade" se observa, por exemplo, quando o professor define o sujeito como "aquele que pratica a ação". Se o aluno objetar com um exemplo como:

(9) Marília foi perseguida por um São Bernardo onde Marília não pratica a ação, o professor poderá re trucar: "Ah, mas aí é voz passiva". Ou seja, a definição de sujeito dada não está completa; na verdade, a definição nunca será completada, de maneira a proporcionar sempre um caminho de fuga a contra-exemplos embaraçosos. As conseqüências da exposição dos alunos a essas situações não podem ser favoráveis nem à sua formação intelectual, nem ao seu respeito pela matéria.

Aquilo que desejamos, como professores, é evidente mente uma boa gramática da língua, não uma explicitação da DGImp. — que, embora quase sempre mais correta do qu a DGEx., está longe de refletir uma análise coerente da estrutura da língua. Mas essa explicitação da DGImp. pode ter valor educativo, porque nos põe em guarda contra situações como a atual: a convivência pacífica com contradições gritantes. Uma formação gramatical intelectualmente sadia só pode ser atingida através de um exame racional e rigoroso do fenômeno da linguagem e da estrutura da língua, nunca através de princípios desconexos e, o que é pior, ministrados dentro de um esquema de autoridade. Mas isso pressupõe a existência de uma teoria gramatical que possa dirigir o esforço de análise e compreensão do funcionamento da língua. E antes de ser possível trabalhar com essa teoria, será preciso desenvolver (nos alunos e em nós mesmos) a atitude intelectual conveniente: exigi remos de nós mesmos um nível de rigor e coerência do qual os estudos gramaticais tradicionais estão muito longe; e nos colocaremos diante do objeto de estudo não como meros receptores de conhecimentos já produzidos, mas como críticos e criadores de conhecimento novo. Nenhuma mudança no conteúdo conceptual da disciplina terá utilidade sem essa mudança de atitude.

Acredito que a tentativa de explicitação das desarmonias entre DGEx. e DGImp., assim como os outros aspectos da crítica à gramática tradicional (ou à gramática não-tradicional, bem entendido!) é um caminho para a

criação de um tipo de atitude que permitirá à nossa disciplina ser realmente útil na formação intelectual dos estudantes.

3 Três problemas básicos

Vamos agora examinar com algum detalhe três tópicos que devem ser esclarecidos e explicitados antes de se empreender qualquer tentativa de renovação da gramática. Esses tópicos são: a relação entre o aspecto semântico e o aspecto formal da linguagem; a definição de "paradigma gramatical"; e a distinção entre classes e funções. Esses três pontos me parecem particularmente relevantes, razão pela qual optei por tratá-los em posição de evidência nas próximas subseções.

O formal e o semântico

Como seria uma "boa" gramática do português? Idealmente, ela deveria desempenhar a contento duas funções: (a) descrever as formas da língua (isto é, sua fonologia, sua morfologia e sua sintaxe); e (b) explicitar o relacionamento dessas formas com o significado que veiculam. Esses estão entre os objetivos de qualquer gramática, ainda que nem sempre estejam claramente formulados ou bem diferenciados. Assim, quando encontramos a afirmação de que a forma reflexiva formada de verbo seguido de pronome oblíquo de pessoa igual à que o verbo se refere" (BECRARA, 1968, p. 127.), podemos entendê-la como a descrição de parte da estruturação formal do português: o fato de que os verbos podem ser seguidos de pronome oblíquo idêntico em pessoa ao sujeito desse verbo. Por outro lado, na passagem [forma reflexiva] indica que a pessoa é, ao mesmo tempo, agente e paciente da ação verbal" (Id., ib id), o autor procura estabelecer um relacionamento entre a forma reflexiva e um significado específico. Naquilo que têm de descritivo (não-normativo), portanto, as gramáticas são tentativas de explicitar esses dois componentes da estrutura da língua: suas formas e o relacionamento dessas formas com os respectivos significados.

Com efeito, simplificando um pouco as coisas, pode-se dizer que natural consiste de um conjunto de recursos formais que servem (juntamente com o contexto extralingüístico) para transmitir um conteúdo. As relações entre a forma e o conteúdo são extremamente complexas, e em grande parte permanecem obscuras ainda hoje para os lingüistas. Qualquer doutrina gramatical que venha a ser proposta precisa levar em conta a complexidade dessa relação, sob pena de cair em inadequações sérias a cada passo. Na presente seção tentarei mostrar que a GT incorre justamente nesse erro; e procurarei sugerir caminhos para evitar esse tipo de incorreção.

Um dos muitos pressupostos não-confessados da GT é o de que a relação entre o aspecto formal e o semântico da linguagem é relativamente simples. Ë algo como acre ditar que para cada forma sintática ou morfológica existe um significado básico e só um (ou uns poucos), de maneira que a explicitação da relação forma—sentido seria na essência uma questão de

justaposição: a forma X tem o significado Y. Encontramos essa crença por trás de definições mistas (formais e semânticas) do seguinte tipo:

(10) "Verbo é (. . .) a palavra que exprime um fato (ação, estado ou fenômeno) representado no tempo. (. . .) O verbo apresenta as variações de número, de pessoa, de modo, de tempo e de voz." (CUNHA, 1975, p. 253.)

Temos aí, para a mesma noção de "verbo", duas definições: uma semântica (a palavra que exprime um fato representado no tempo), outra formal (a palavra que apresenta variações de número, pessoa, modo, tempo e voz). Por trás dessa dupla definição, naturalmente, há o pressuposto de que qualquer palavra que corresponda à primeira parte da definição também corresponderá à segunda — ou seja, de que a relação entre as propriedades semânticas e as formais do verbo é simples e direta, podendo ser expressa por uma mera justaposição de definições, como está em (10).

A verdade, entretanto, é que isso não ocorre'. É bastante fácil encontrar palavras que correspondam a uma das definições e não à outra. Por exemplo:

Aqui temos chuva, palavra que exprime um fenômeno, e este está representado no tempo (ontem); não creio, porém, que seja interessante classificá-la como verbo. E, certamente, chuva não apresenta o conjunto de variações apontado como característica (formal) dos verbos.

Uma resposta possível a essa objeção é a de que em

(11) o tempo não está expresso pela palavra chuva, através de um morfema preso, como em:

(12) choveu ontem e meu canteiro ficou estragado

Choveu seria um verbo porque o tempo está representado morfologicamente, dentro da própria palavra (choveu, chove, choverá. .). Dessa forma, "representado no tempo" teria de ser entendido como referindo-se apenas a uma marca morfológica: o elemento marcador de tempo deve estar incluído na própria palavra em questão, para que esta seja classificada como verbo. Mas esse refinamento da definição semântica (introduzindo, na verdade, um elemento formal) nos deixaria em dificuldades quando considerássemos casos como:

(13) gato come rato

A palavra come, que exprime um fato (uma ação), e que é usualmente considerada um verbo, não está "representada no tempo", pelo menos se consideramos "tempo" no seu sentido habitual. Com efeito, (13) exprime uma afirmação geral, intemporal, sobre gatos (e ratos). Essa frase se coloca na classe das afirmações universais do tipo: O homem é mortal, A água ferve a cem graus etc., das quais não me parece que se possa dizer que estejam "representadas no tempo". Segundo a definição semântica proposta, mesmo com a restrição de representação no tempo apenas à representação morfológica, acabamos sendo obrigados a considerar come em (13) como um não-verbo — decisão que certamente não agrada a ninguém.

Se continuássemos refinando a definição semântica, poderíamos chegar a uma formulação satisfatória do que significa exatamente

"representado no tempo". Mas no momento é mais importante reconhecer precisamente o que é que estamos tentando fazer ao explicitarmos essa noção: na verdade, procuramos uma relação entre o tempo gramatical, formal, que é aquilo que opõe a forma come às formas comia, comeu etc. e uma interpretação semântica (uma referência temporal) que seja própria ao elemento formal a que chamamos "tempo". Ora, o tempo gramatical não é simplesmente uma representação formal do tempo cronológico; aliás, o fato de que damos a ambos a mesma designação ("tempo") não nos deve enganar quanto a suas naturezas profundamente diferentes. Já vimos que em (13) o presente do indicativo exprime um fato intemporal. Em:

(14) pode deixar que eu frito os bolinhos

o mesmo tempo verbal se interpreta como um tempo (referencial) futuro. E em:

(15) nesse momento, D. Pedro tira a espada e grita:

"1 ou M!"

o presente formal denota um passado referencial. Temos aqui um exemplo bem claro da complexidade da relação forma—significado. É, evidentemente, necessário explicitar essa relação, que existe e é a razão de ser principal da própria linguagem. Mas não se espere atingir essa explicitação sem lançar mão de uma teoria semântica muito mais rica, complexa e estruturada do que o conjunto de noções desconexas e improvisadas que passa por semântica na GT.

Resta considerar a definição formal de "verbo". Podemos formulá-la assim:

(16) Verbo é a palavra que pertence a um paradigma cujos membros se opõem quanto a número, pessoa e tempo 2

Aqui, evidentemente, "número", "pessoa" e "tempo" se referem a morfemas flexionais específicos, e não a categorias de significado.

Com uma definição formal como (16) será possível evitar os sérios problemas que acabamos de ver, e que são típicos de definições semânticas. Em geral as definições formais são mais fáceis de elaborar, de testar e de relacionar entre elas; isso porque, creio, existem hoje teorias sintáticas muito mais desenvolvidas e precisas do que as teorias semânticas; ou ainda, dito de outro modo, o fenômeno sintático é muito melhor compreendido do que o fenômeno semântico. Note-se que isso não quer dizer que a descrição semântica não seja importante; significa apenas que, no momento, qualquer descrição semântica de uma língua deverá ser menos completa e satisfatória do que as descrições formais (sintáticas e morfológicas) que estamos em condições de produzir.

O critério que nos permite identificar tão facilmente correu como um verbo, e corrida como um não-verbo, não é o significado de cada uma dessas palavras, mas o fato de que correu se liga a um paradigma que inclui corri, correram, corre (formas que se opõem quanto à pessoa, ao tempo e ao número), ao passo que corrida pertence a outro paradigma, que inclui, por exemplo, corridas e corridinha, que se opõem quanto ao número (mas um

tipo de “número” morfologicamente diferente do número verbal) e quanto ao grau.

(16) seria, provavelmente, a definição de “verbo” que se encontra na DGImp., e que realmente aplicamos ao classificarmos palavras como verbos ou não-verbos. É uma definição muito mais fácil de aprender e de aplicar do que a definição semântica (muito embora não deixe de lançar mão de noções subjacentes de certa complexidade; ver a subseção seguinte). Raramente se observam hesitações na aplicação dessa definição, e o grau de semelhança formal entre as palavras a que chamamos “verbos” é muito grande.

No entanto, antes de nos sentirmos autorizados a incluir definições do tipo de (16) na gramática, será preciso discutir e esclarecer a noção de “paradigma gramatical”, de que (16) faz uso de modo crucial. Essa é outra das noções básicas subjacentes à prática da análise gramatical, e que nunca são explicitamente consideradas. Nas próximas páginas, portanto, vou examinar a noção de “paradigma gramatical”.

A noção de “paradigma gramatical”

Vou inicialmente esclarecer uma questão de nomenclatura: estou tomando aqui o termo “palavra” em um sentido talvez mais restrito do que o usual. Assim, para mim, homem e homens são duas palavras distintas, ainda que membros de um mesmo paradigma, e não duas formas da mesma palavra. As razões para isso serão dadas no final da presente subseção; por ora, tenhamos em mente que quando a GT fala de “várias formas da mesma palavra” (por exemplo, um singular e seu plural), eu prefiro dizer “várias palavras pertencentes ao mesmo paradigma”.

Passemos agora ao problema da conceituação de “paradigma”. Uma das maneiras de classificar as palavras formalmente é considerar suas possibilidades de variação

morfológica; assim, um verbo varia em pessoa (isto é, pertence a um paradigma cujos membros diferem quanto à pessoa); um adjetivo varia em gênero, um pronome pessoal em caso. Esse tipo de classificação é tradicional, e provavelmente conveniente em muitos casos. Mas não é isento de problemas: na verdade, creio que se pode mostrar que em certos casos uma classificação puramente morfológica é, estritamente falando, impossível. Como conseqüência, a classificação de base sintática tem um caráter fundamental.

Tomemos o caso dos adjetivos: podemos defini-los como a classe das palavras que variam em gênero, número e “grau” (normal X superlativo). Branco é então um adjetivo porque existem branco, branca, branquíssimo, brancos etc. Brancura e branqueamos não são adjetivos, pois não variam da mesma forma. Isso parece evidente e difícil de contestar, mas só será sustentável se conseguirmos responder à seguinte pergunta: Por que consideramos branco, branca, branquíssimo e brancos como pertencentes a um paradigma (“o adjetivo branco”), branqueamos e branqueio a outro (“o verbo branquear”), e ainda brancura, brancuras a um terceiro (“o substantivo brancura”)? Não vejo possibilidade de responder a essa pergunta em termos morfológicos sem cair em circularidade. E uma resposta baseada

na semântica incorreria nas dificuldades usuais desse tipo de definição: poderíamos dizer que branco, branquíssimo etc. "exprimem qualidade", e que por isso devem ficar juntos em um paradigma. Mas que dizer de "os topázios brilham muito", em que brilham exprime uma qualidade? Por aqui, é fácil ver, não há saída visível.

Se quisermos salvar a caracterização morfológica dos adjetivos, então, teremos de justificar a inclusão de branca, branco, brancos e branquíssimo em um paradigma único. Obviamente, isso deve ser feito sem mencionar o relacionamento morfológico como critério: ou seja, sem chegar a uma justificativa do tipo "essas formas pertencem ao mesmo paradigma porque são variantes morfológicas da mesma palavra", o que é tautológico. Será preciso procurar outros critérios para capturar a noção tradicional de "paradigma". Como se trata de uma noção útil em gramática, vale a pena determo-nos um pouco tentando conceituá-la com, alguma clareza.

Para chegar a essa conceituação, partamos da condição seguinte: todos os membros de um paradigma devem ter pelo menos um morfema em comum. Acho que essa condição captura uma parte da noção tradicional; ficam assim excluídos casos como o de casa e sempre, que decididamente não podem pertencer ao mesmo paradigma.

A primeira condição estipulada, embora necessária, não é suficiente. Se nos limitarmos a ela, teremos de colocar no mesmo paradigma formas como correm e corrida, ou ainda falei e comprei, pois em ambos os casos há morfemas em comum: no primeiro caso, o radical corr-; no segundo, o sufixo de modo-tempo-pessoa -ei. Logo, necessitamos de mais alguma condição para refinar a definição de "paradigma".

Os exemplos citados no parágrafo precedente sugerem a introdução de uma condição nos seguintes termos: todos os membros de um paradigma devem pertencer à mesma classe de palavras. Desse modo poderemos, pelo menos, separar correm de corrida em paradigmas distintos. Aqui, novamente, há o perigo da tautologia: será necessário definir as classes de palavras em termos não-morfológicos, já que qualquer definição morfológica depende da noção de "paradigma", que estamos tentando elucidar. As classes de palavras, como veremos detalhadamente mais tarde, serão definidas segundo critérios sintáticos: pertencem a uma mesma classe palavras que ocorrem no mesmo conjunto característico de ambientes sintáticos. Por exemplo, branco, branca e branquíssima pertencem à mesma classe por ocorrerem nos mesmos ambientes sintáticos. Um desses ambientes pode ser expresso assim: ocorrência logo após uma seqüência de artigo + substantivo, formando (os três) um sintagma nominal:

(17) comprei um cachorro branco/a galinha branca/ uma camisa branquíssima

Já *brancura* ou *branqueia* não têm essa propriedade sintática, e portanto não se colocam na mesma classe de branco.

Como veremos, a condição de que os membros de um paradigma pertençam à mesma classe de palavras ainda nos poderá trazer problemas, se quisermos realmente respeitar a distribuição tradicional das palavras em

paradigmas. Por ora, vamos adotá-la e ver quais são as conseqüências imediatas.

Agora já podemos incluir em um paradigma único formas como *correm, corremos, corria, correrei* (que recebem tradicionalmente o nome coletivo de "verbo *correr*", isto é, o paradigma de *correr*); ficam excluídas desse paradigma formas corno *corrida* ou *corridinha,* que por sua vez integram um outro paradigma. A justificação é que, além de terem um morfema em comum, correm, corremos etc. pertencem à mesma classe ("verbos"); o mesmo se passa com *corrida, corridinha,* que são "substantivos".

No entanto, a definição ainda não se pode considerar completa, porque nada nos impede de colocar *falei* e *comprei* em um mesmo paradigma. A tradição não autorizaria dizer que *falei* e *comprei* são formas da mesma palavra; vamos aceitar como correta essa posição, pelo menos para efeitos da presente discussão. Assim, concluiremos que ainda falta alguma coisa na definição de "paradigma".

Talvez seja indicado incluir a condição de que os membros de um paradigma só possam diferir quanto a seus morfemas *flexionais*; os morfemas não-flexionais (radicais e morfemas derivacionais) são comuns a todos os membros de um paradigma. Isso nos possibilita separar *falei* e *comprei* em paradigmas distintos (pois diferem quanto ao radical); faremos o mesmo com *fazer, refazer e desfazer* (pois diferem quanto a morfemas derivacionais). Dessa maneira, para capturar as distribuições tradicionalmente feitas, definiremos "paradigma" do seguinte modo (conforme veremos, a definição ainda terá de ser melhorada):

(18) Paradigma é um conjunto de palavras que pertencem à mesma classe e que diferem apenas quanto a morfemas flexionais.

Essa definição, é claro, depende de podermos distinguir claramente morfemas flexionais de morfemas não-flexionais (em especial, de morfemas derivacionais), o que pode não ser muito fácil. Aronoff comenta que a distinção entre fenômenos derivacionais e flexionais

"é delicada, e às vezes fugidia, mas não obstante importante'. (ARONOFF, 1976, p. 2.)

Não conheço definição inteiramente satisfatória de derivação e flexão (ver proposta de NIDA, 1949, com crítica, no trecho citado de Aronoff). Mas é possível, desde já, atingir uma conceituação aproximada que poderá ser útil. Vou discutir brevemente este ponto, importante para nós, porque dele depende uma noção clara de "paradigma".

Aronoff parece considerar como uma característica básica da flexão, frente à derivação, a regularidade de ocorrência das diversas formas comparáveis. Assim, parece-me, ele chamaria de flexional uma variação como *correm, corro, corremos, correndo, correr* porque para praticamente todas as formas que terminam com o morfema *-r* 'infinitivo' existem outras formas com os morfemas (ou seqüências de morfemas, pouco importa) *-m, -o, -mos,-ndo*. Podemos acrescentar que a esse paralelismo formal corresponde um paralelismo semântico bem claro. Já no caso de *fazer* e *desfazer* não encontramos essa regularidade; existe *atar* e *desatar*, *montar* e *desmontar*,

mas não existe forma com *des*- para riscar (o antônimo é apagar), acender (usa-se novamente apagar), abrir (fechar) etc. É a regularidade básica da flexão que nos autoriza a falar de "supletivismo" nos raros casos excepcionais, como *ser, fui,* considerado paralelo a *cantar, cantei,* apesar da diferença entre as duas formas. Admitimos, *a priori*, que todo verbo precisa ter um pretérito perfeito, mas nem todo verbo precisa ter uma forma antônima em *des*-; daí não dizermos que apagar é forma supletiva do antônimo de acender (substituindo um inexistente desacender).

Esse critério de distinção entre flexão e derivação não é tão nítido quanto desejaríamos, e parece mais uma questão de grau: depende, em última análise, da quantidade de exceções em cada caso. Mas é o que temos à mão no momento, e o que vou adotar (não lançarei sobre Aronoff a responsabilidade da interpretação acima, que não está explícita em seu trabalho). O critério basta para decidir a maior parte dos casos, e vou, portanto, mantê-lo como o que melhor exprime o uso implícito dessa noção por parte dos autores.

Podemos agora retomar a definição (18), para tentar aperfeiçoá-la. Conforme nos lembramos, (18) estabelece duas condições para que duas ou mais palavras integrem um mesmo paradigma: pertencerem à mesma classe e serem idênticas no que diz respeito aos morfemas não-flexionais (mesmo radical, mesmos morfemas derivacionais). Ou seja, ficam definidas quatro possibilidades de relação entre as palavras, a saber:

A: mesma classe, identidade de morfemas não-flexionais;

B: classes diferentes, diferença de morfemas não-flexionais;

C: mesma classe, diferença de morfemas não-flexionais;

D: classes diferentes, identidade de morfemas não flexionais.

Para cada uma dessas possibilidades, examinaremos um exemplo, a fim de apurar como funciona o agrupa mento tradicional, e verificar se essas duas condições são suficientes, e se são ambas necessárias para caracterizar a noção de "paradigma", tal como utilizada implicitamente na GT.

A possibilidade *A* é exemplificada, por exemplo, pelas palavras *branco* e *branca*: ambas são adjetivos, e sua única diferença reside nos morfemas flexionais *-o/-a*. Sabemos que a GT coloca essas duas palavras no mesmo paradigma (ou, o que sabemos ser equivalente, considera-as duas formas da mesma palavra). A possibilidade *B* se verifica com o par *branco/brancura*: pertencem a classes diferentes, e além disso diferem quanto ao morfema derivacional *-ur*(a). A solução tradicional é colocar as palavras em paradigmas diferentes: *brancura* não é chamada de uma outra forma da palavra *branco*. Observamos a possibilidade *C* no par *fazer/desfazer*: pertencem ambas à classe dos verbos, mas diferem quanto ao morfema derivacional des-. A GT os separa em paradigmas distintos.

Até agora, portanto, parece que a exigência é de que duas palavras precisam satisfazer ambas as condições especificadas em (18) para que possam ser colocadas no mesmo paradigma. Mas a possibilidade *D* desmente isso. Na verdade, é difícil encontrar pares de palavras que sejam idênticas

quanto a seus morfemas não-flexionais (isto é, que só tenham diferenças flexionais) e que não pertençam à mesma classe. Mas esses casos existem, e, quando ocorrem, a GT coloca as palavras em questão *no mesmo paradigma*, ao contrário do que seria de esperar.

É o caso das formas "nominais" do verbo, como os infinitivos, por exemplo. Essas formas se relacionam de maneira regular com os verbos; para cada *corro* existe um *correr* etc. Entretanto, há muitas razões para se acreditar que os infinitivos não são, sintaticamente, da classe dos verbos. Seu comportamento gramatical é predominantemente nominal, o que é aliás geralmente reconhecido na GT. Celso Cunha, que define, como vimos, o verbo como incluindo "representação no tempo", afirma do infinitivo (e das outras formas nominais) que:

"identificam-se pelo fato de não poderem exprimir por si nem o tempo nem o modo". (CUNHA, 1975, p. 456.)

Dessa maneira, é provável que os infinitivos não devam ser colocados na classe dos "verbos", que inclui as formas não-nominais. Mas a regularidade de sua relação com essas formas faz com que a diferença deva ser considerada como de flexão, e não de derivação. Em conseqüência, temos aqui a situação descrita em *D*: *corro* e *correr* são idênticos quanto aos morfemas não-flexionais, mas pertencem a classes diferentes. E nesse caso, conforme sabemos, a GT os coloca no mesmo paradigma: *corro* e *correr* são considerados formas da mesma palavra. Esse critério, se aplicado consistentemente, significa que a explicitação do uso tradicional da noção de "paradigma" não deve conter a condição de que as palavras interessadas pertençam à mesma classe. A definição se reduz a:

(19) Paradigma é um conjunto de palavras que diferem apenas quanto a morfemas flexionais.

Sabemos que a distinção entre morfemas flexionais e não-flexionais depende da regularidade da associação dos afixos com os radicais. Temos aqui um ponto que deverá ser mais cuidadosamente investigado: Como definiremos essa regularidade? Como trataremos casos como *ser/fui*, onde a correspondência morfológica é regular, mas a representação fonológica não é? Muitas questões importantes como essa terão de ficar à espera de investigação futura.

Resta um ponto a esclarecer. A nomenclatura tradicional, como vimos, não costuma falar de paradigmas, mas de palavras; isto é, *corro, correr e corríamos* não são chamados membros de um paradigma único, mas "formas da mesma palavra". A razão por que não adoto essa nomenclatura, preferindo referir-me a corro e correr como duas palavras distintas, é que, se definirmos como "uma palavra" uma entidade como o "verbo *correr*" da GT, essa entidade será sintaticamente incoerente. Seremos obrigados a dizer que uma palavra pode ter comportamento sintático diverso (e, conseqüentemente, pertencer a mais de uma classe) segundo a sua flexão: um infinitivo não se pode classificar da mesma forma que um presente do indicativo. Isso ocorre porque a noção de "paradigma" é morfológica, e não sintática. Acho que essa posição leva a situações con fusas, e, portanto, opto

por utilizar o termo “palavra” no sentido restrito que considera corro uma palavra e correr outra palavra distinta.

Por último, pode-se perguntar qual é, exatamente, a utilidade da noção de “paradigma” na descrição gramatical. Essa noção, antes de mais nada, permite a formulação de definições morfológicas como a de que o verbo é a palavra que “varia em pessoa, tempo” etc. — ou seja, o verbo é a palavra que pertence a um paradigma cujos membros diferem quanto a essas categorias. Já mencionei acima a conveniência de tais definições. Por outro lado, os paradigmas possibilitam a apresentação compacta de grupos de palavras flexionalmente relacionadas, como por exemplo os paradigmas verbais, o que é vantajoso do ponto de vista didático. Finalmente, parece-me que os paradigmas terão importância quando da descrição da estrutura semântica da língua, pois o relacionamento semântico entre os membros de um paradigma é sistemático.

Classes e funções

Passemos agora ao terceiro dos três problemas básicos a serem discutidos, o da distinção entre “classe” e “função”.

A GT faz uso de certo número de noções, nunca explicitadas; conseqüentemente, faz muitas vezes um uso incoerente dessas noções, ou deixa de aproveitar as vantagens práticas e teóricas de sua existência. É esse o caso, acredito, das duas importantes noções de “classe” e “função”. Embora ambas façam parte da parafernália teórica da GT, a distinção entre elas é pouco satisfatória.

A noção de “classe” encontra-se reconhecida, ainda que não bem definida, nas gramáticas, mas seu uso é pouco sistemático. Admite-se sempre a necessidade de classificar as palavras, e a doutrina fornece nomes para essas classes (“verbos”, “advérbios”, “pronomes” etc.). Além dessas classes, existem outras, que não são explicitamente reconhecidas como tais, mas que também recebem nomes: termos como “oração”, “frase”, “oração subordinada” se referem na verdade a classes de formas, ou a suas sub-classes. E, como quaisquer outras classes, podem ser definidas pela sua distribuição sintática, sua estrutura interna, ou (com as limitações que conhecemos) suas propriedades semânticas.

No entanto, nem todas as classes são explicitadas. Vejamos o caso de uma classe extremamente freqüente e importante, mas que não é em geral reconhecida pela GT (salvo raras exceções): a dos sintagmas nominais. Por que seria importante incluir explicitamente essa classe na gramática?

Formas como as exemplificadas abaixo:

(20) Carminha

a Carminha

aquela moça do terceiro andar

uma funcionária da Universidade que eu conhecia

apesar de suas grandes diferenças estruturais, têm em comum traços sintáticos muito importantes: por exemplo, todas elas podem ser sujeito de

uma oração; todas podem ser objeto direto; todas podem vir precedidas de preposição, em funções tais como adjunto adnominal ou objeto indireto. Por outro lado, nenhuma dessas formas pode ser o núcleo de um predicado verbal, nem aparecer coordenada com a conjunção e mais um adjetivo. Em outras palavras, as quatro formas de (20) têm comportamento sintático semelhante. Ora, uma das funções essenciais das classes de formas (por exemplo, das classes de palavras) é justamente permitir a descrição compacta do comportamento sintático das formas. As quatro formas de (20) deveriam, pois, ser colocadas em uma classe, o que a GT não faz; não existe sequer um termo tradicional para essa classe. Aqui utilizaremos "sintagma nominal", designação consagrada em Lingüística.

A inexistência de um termo para designar comodamente o sintagma nominal (ou melhor, a inexistência de uma noção clara de sintagma nominal) obriga as gramáticas a descreverem o comportamento sintático dessa classe de maneira desnecessariamente complicada e sem unidade em outras palavras, de maneira não-sistemática. Assim, Cunha (1975) diz que um sujeito pode ter como núcleo: um pronome pessoal; um substantivo; um pronome demonstrativo; um pronome relativo; um pronome interrogativo; um pronome indefinido; um numeral; uma palavra ou expressão substantivada; uma oração substantiva; mais de um substantivo; mais de um pronome; mais de um numeral; mais de uma palavra ou expressão substantivada; mais de uma oração substantiva; ou "outras combinações" — ao todo, quinze possibilidades de sujeitos claros (Ou mais, já que algumas possibilidades estão abreviadas). Essa forma de descrever os fatos, além de seu caráter obviamente pouco compacto, apresenta dois defeitos graves: primeiro, deixa de explicitar a estrutura propriamente dita dos sujeitos, dando apenas a lista dos possíveis *núcleos* do sujeito. Assim, *Carminha*, *a Carminha* ou *aquela moça do 3.° andar* caem no mesmo caso, pois em todos os exemplos o núcleo é um substantivo.

O segundo defeito, mais grave, desse tipo de descrição é que ele nos obrigará a repetir a *mesma* lista de quinze possibilidades quando formos explicitar quais são os possíveis núcleos de um objeto direto, ou de um objeto indireto, ou de um adjunto adnominal. Como se vê, estamos deixando escapar uma generalização importante da língua; estamos tratando como simples coincidência um fato estrutural de importância, a saber, o de que a mesma classe de formas (o sintagma nominal) pode aparecer em todas essas funções sintáticas.

Estabelecendo o sintagma nominal como uma classe de formas, poderemos solucionar essas deficiências de modo bastante simples. A estrutura interna do sintagma nominal será definida uma só vez na gramática: o sintagma nominal se compõe de um substantivo, ou de artigo seguido de substantivo, ou de pronome pessoal etc. (a composição do sintagma nominal é bem complexa). A partir daí, diremos que o sujeito é sempre composto de um sintagma nominal; o mesmo afirmaremos do objeto direto; o objeto indireto se compõe de preposição mais sintagma nominal (excetuados os casos de clíticos como *lhe*, que precisarão ser tratados à parte, dado seu comportamento sintático peculiar). Trata-se, simplesmente, de estender sistematicamente às seqüências de palavras o tratamento que se

dá às palavras individuais. Não costumamos dizer que as palavras *correr, dizer, mandar, ir* e mais outras milhares variam em pessoa, mas antes que o verbo varia em pessoa; igualmente, diremos que um sintagma nominal pode ser sujeito — e não que um substantivo, uma seqüência de artigo mais substantivo, um pronome pessoal etc. podem ser sujeitos. E também com isso capturaremos na gramática o fato de que as formas que podem desempenhar a função de sujeito são as mesmas que podem desempenhar a de objeto direto.

Esse exemplo pode ser útil para deixar clara a distinção entre “função” e “classe”. “Sujeito” é uma função, isto é, um dos aspectos da organização formal da oração. Uma função sintática se define através de relações sintagmáticas entre os diversos termos da oração: ordem das palavras, concordância, regência etc. Assim, a função de sujeito se caracteriza por certas posições na oração, e por estar em relação de concordância de pessoa e número com o verbo.

A partir da função de sujeito poderíamos definir uma classe de formas: pertenceriam a essa classe todas as formas que pudessem ter a função de sujeito. Mas imediatamente se apresenta outro fato importante: a classe das formas que podem ser sujeitos é idêntica à das formas que podem ser objetos diretos; precedidas de preposição, essas mesmas formas podem ser adjuntos adnominais etc. Ou seja, a classe não se identifica com a função, nem tem uma vinculação biunívoca com ela: uma classe se define pela relação paradigmática entre as diversas formas que podem desempenhar uma (ou várias) funções. Diremos, então, que existe a *classe* dos “sintagmas nominais”, e que um sintagma nominal pode desempenhar várias funções sintáticas.

É necessário distinguir “classe” de “função” não apenas para possibilitar uma descrição clara do fato de que várias funções podem ser desempenhadas por uma só classe, mas também porque essas duas noções são distintas em princípio, e exprimem aspectos muito diferentes do fenômeno lingüístico. Uma classe é um conjunto (não necessariamente finito) de formas lingüísticas; uma função é um princípio organizacional da linguagem. Assim, quando digo que uma forma qualquer é um sujeito, não a estou colocando em uma classe (ainda que sua classe possa ser inferida do fato de ela ser sujeito), mas estou antes afirmando que, naquela sentença, a forma preenche um dos lugares que devem ser preenchidos para que se tenha uma sentença gramatical da língua.

Colocadas as coisas dessa maneira, verifica-se que a nomenclatura tradicional é especialmente deficiente no que diz respeito a nomes de classes; correspondentemente, muitas classes não são devidamente reconhecidas na gramática. A única preocupação consistente que encontramos na GT é a de classificar as palavras, de modo que nenhuma fique isolada (ainda que seja posta entre as “de difícil classificação”...). Mas essa preocupação não se estende às seqüências de duas ou mais palavras. Isso faz com que se percam generalizações como as que vimos acima para o caso do sintagma nominal.

Pode-se multiplicar a exemplificação de casos de classes importantes gramaticalmente, mas não reconhecidas na GT; por falta de espaço, não poderemos prosseguir no exame de tais casos, além do já visto do sintagma nominal. Mas certamente isso terá de ser feito quando da elaboração da nova gramática portuguesa. Aqui nos limitaremos a algumas considerações finais.

Ao se considerar uma dada estrutura, é necessário não apenas perguntar qual é a função de cada constituinte, mas ainda a que classe pertence cada constituinte. Não é o bastante saber, por exemplo, que as orações em português podem ter sujeito; é preciso saber também que essa função é ocupada por uma classe particular de formas (sintagmas nominais), e que essa mesma classe pode ocupar outras funções (objeto direto etc.). A classificação das palavras é apenas um caso particular da classificação das formas sintáticas, e desprezar esse fato equivale a deixar de exprimir muitas generalizações importantes sobre a estrutura da língua.

4 As bases da nova gramática

Problemas da descrição semântica

Até o momento, a ênfase da discussão tem incidido sobre uma crítica às posições tradicionais e sobre a tentativa de explicitar alguns de seus pressupostos. Vamos agora passar a considerar algumas questões (relacionadas ou não com as questões já discutidas) na perspectiva da construção da nova gramática Tentarei dar, nas seções que se seguem, em uma série de “quadros”, uma proposta para uma gramática mais adequada do português. Se a imagem aqui mostrada parecer muito preliminar e fragmentária, isso significa apenas que há muito o que fazer neste campo. Uma estagnação de várias décadas não se vence em pouco tempo, nem com pouco esforço; aqui vou tentar, no máxi mo, definir alguns problemas e apontar caminhos. A seleção de temas é algo arbitrária: escolhi aqueles nos quais já me detive, e sobre os quais tenho mais a dizer. Restam grandes áreas a investigar, mesmo no que concerne aos fundamentos da nova linguagem gramatical; fica aqui, novamente, o convite à pesquisa.

Já dei anteriormente minha opinião sobre o papel da descrição semântica em uma gramática. Na ocasião, afirmei que a descrição semântica é importante e altamente interessante, mas apresenta uma série de dificuldades, oriundas em última análise da ignorância relativa em que nos encontramos dos fenômenos semânticos em geral (digo “ignorância relativa” considerando a comparação com a sintaxe, a fonologia etc.). No que se segue, entrarei em alguns pormenores, a fim de esclarecer melhor a natureza e as dimensões do problema.

Semântica e pragmática

Um dos problemas que se colocam para qualquer estudo que leve em conta o significado das formas lingüísticas é que não é claro se essas formas

realmente possuem um significado independente do contexto situacional em que são utilizadas.

Em muitos casos esse problema não é imediatamente evidente. Tomemos uma configuração formal determinada: as sentenças que se caracterizam, na escrita, pela presença de um ponto de interrogação e pela ausência de palavra interrogativa como *quem, onde, quando* etc. Chamarei a essas sentenças "interrogativas sim/não", para distingui-las das "interrogativas Q", que se iniciam por palavra interrogativa. Tenhamos em mente sempre que o termo "interrogativa" designa uma configuração formal, e não um tipo de interpretação semântica; ou seja, é um termo exclusivamente sintático.

Em muitos casos a interrogativa sim/não se interpreta inequivocamente como contendo o ingrediente semântico "solicitação de valor de verdade". Um exemplo é a frase:

(21) você sabe falar javanês?

Essa frase é (praticamente) sempre entendida como equivalente a "Diga-me se a afirmação 'você sabe falar javanês' é verdadeira ou falsa"; por isso disse eu acima que (21) contém o ingrediente "solicitação de valor de verdade" (abreviadamente, SVV) em sua interpretação semântica. Somos tentados, então, a afirmar que a configuração formal exemplificada em (21), isto é, a interrogativa sim/não, só pode ser interpretada como SVV. Essa crença se encontra em geral nas gramáticas. (Ver, por exemplo, QUIRK et al., 1972, p. 387; CUNHA, 1975, p. 178-9.)

No entanto, no caso das interrogativas sim/não, como em muitos outros, a relação entre forma e significado é complexa. Vejamos outra frase, à primeira vista idêntica a (21) no que nos interessa:

(22) você sabe onde fica a Praça 7?

Poderíamos analisar (22) de maneira semelhante a (21), e chegaríamos à conclusão de que (22) é também uma SVV: o falante pede ao interlocutor que lhe diga se sabe ou não onde é a Praça 7. Nesse caso, uma boa resposta para (22) seria *sim*, que aliás é uma boa resposta para (21). Mas isso obscurece um fato importante, aplicável a (22) mas não a (21) — ou, pelo menos, muito mais dificilmente aplicável a (21) — a sentença (22), pode ser utilizada, sem nenhuma modificação formal, para solicitar de alguém que nos diga onde é a Praça 7. Nessa acepção, que é provavelmente a mais natural para (22), a interpretação semântica não contém o ingrediente SVV; conseqüentemente, uma resposta *sim* já não bastaria. Uma boa resposta poderia ser, por exemplo: "Siga em frente mais dois quarteirões".

Parece, pois, que temos duas "camadas" de significado: o significado "literal" (segundo o qual (21) e (22) são paralelas, tendo ambas a acepção SVV), que não de pende da situação concreta em que a frase é enunciada; e o significado "final", que depende, para sua depreensão, não apenas da análise da expressão lingüística, mas ainda de se levarem em conta fatores extralingüísticos, tais como a situação de comunicação. Por exemplo, a frase (22), dita por uma pessoa que aborda um estranho na rua, dificilmente seria interpretada como SVV. Mas se fosse dita por alguém que está mandando

um empregado ir à cidade fazer um serviço qualquer, já poderia ser interpretada como SVV, e respondida de acordo.

O problema está em que não é claro se realmente há significados literais independentes do contexto. Há duas maneiras de encarar a relação entre o significado literal e o significado final de uma sentença: (a) podemos admitir que o significado literal é básico, e que o significado final é computado a partir do literal, *mais* certos traços do contexto extralingüístico. Ou então: (b) podemos negar a existência de significados literais em oposição a finais, isto é, negar que uma expressão lingüística possa ter significado próprio independentemente do contexto. Segundo esta última posição, toda e qualquer interpretação de uma sentença teria necessariamente de levar em conta o contexto extralingüístico; e, além do mais, o processo de interpretação não incluiria um estágio intermediário que consistiria em computar o significado literal a partir da expressão lingüística "pura".

Para os defensores desta segunda teoria, uma frase como (22), dita na situação em que se solicita informação sobre onde fica a Praça 7, não tem, em nível nenhum de análise, um significado literal diferente do significado final observado; em nenhum nível de análise a interpretação de (22), enunciada naquela situação, teria um ingrediente SVV. Já os defensores da teoria (a) sustentariam que, mesmo enunciada na situação descrita, (22) teria um significado básico de SVV; este seria modificado pela ação de traços do contexto, de maneira a tornar-se um pedido de informação sobre a localização da praça.

Muito embora essa seja uma questão extremamente controversa, não existindo (em minha opinião) uma solução satisfatória, teremos de nos definir por uma das posições, ainda que provisoriamente, meramente como uma hipótese de trabalho: temos uma tarefa a executar, e infelizmente não podemos ir resolvendo todos os grandes problemas teóricos que se apresentam em nosso caminho. Vou optar pela teoria (a), ou seja, admitirei a existência de significados literais, que podem diferir dos significados finais das sentenças. Isso não implica, é claro, negar a relevância do contexto para descrever a interpretação final das expressões lingüísticas; mas autoriza-nos a falar de significado (em certo nível) independentemente do contexto extralingüístico. Essa posição, creio, facilita o trabalho, considerando-se o nível de detalhamento a que precisa chegar uma gramática pedagógica. E, como veremos abaixo, existe alguma motivação adicional para essa decisão, embora seja apenas circunstancial, e não conclusiva.

O estudo do significado literal é geralmente chamado "semântica"; o estudo dos fatores que modificam esse significado literal (para aqueles que colocam assim a questão) é colocado na "pragmática". Desse modo, o dilema que vimos acima, de escolha entre a teoria (a) e a (b), é usualmente formulado em termos de existência ou não de distinção teórica entre a semântica e a pragmática.

Não existe critério universalmente aceito para distinguir pragmática de semântica; e nenhum dos critérios correntes é realmente preciso. Na necessidade de basear-me em alguma coisa, vou ater-me a uma definição que corresponde aproximadamente à idéia implícita que muitos lingüistas

têm dessa distinção (pelo menos quando não estão tentando explicitá-la). Segundo essa definição, a semântica seria o estudo do significado das formas lingüísticas "maximamente descontextualizadas", e a pragmática, o estudo do significado das formas lingüísticas "quando enunciadas em classes particulares de contextos" (citações de LYONS,1977, p. 591).

Essa definição tem seus inconvenientes, O mais sério é o fato de lançar mão da noção de forma lingüística "maximamente descontextualizada". Conforme aponta o próprio Lyons,

"não se pode passar das sentenças às proposições expressas por elas (. .) sem levar em conta certos traços contextuais". (ibid.)

Isto é, a descontextualização total não é possível,, daí sermos obrigados a falar em descontextualização máxima (o que implica graus de descontextualização).

Já vimos que há dúvidas sobre se se pode realmente ter acesso a uma forma lingüística descontextualizada; por definição, toda forma lingüística é enunciada em um contexto. Assim, é melhor encarar a descontextualização não como uma situação absoluta em que se possam encontrar as expressões, mas antes como um eixo cujas extremidades podem ser aproximadas, nunca atingidas. Para esclarecer melhor o que quero dizer, vejamos uma sentença apresentada por escrito, sem contexto lingüístico precedente:

(23) a porta está aberta

Nessa situação, a sentença será interpretada como uma asserção sobre o estado da porta. Compare-se a enunciação da mesma sentença dirigida a alguém que aparece na porta do meu gabinete, parando ali como que hesitando se deve entrar. Nesse contexto, (23) será certa mente interpretada como um convite a entrar, e não como asserção quanto ao estado da porta. Diríamos, então, que (23) está menos contextualizada no primeiro caso do que no segundo.

Podem-se formular perguntas muito embaraçosas quanto a esse método intuitivo de apurar o sentido descontextualizado (literal) das formas lingüísticas. Como sabemos que o primeiro contexto (apresentação por escrito) é verdadeiramente menos contextualizado do que o segundo? Não vejo modo de resolver satisfatoriamente esse problema; no máximo, posso dar alguma evidência (como disse, circunstancial) de que a posição adotada aqui é correta: talvez os falantes tenham algum acesso intuitivo ao significado literal das formas lingüísticas. A própria existência do termo tradicional "sentido figurado" pode ser tomada como evidência disso. Freqüentemente as pessoas concordam quanto ao *status* figurado ou não de determinada forma em determinado contexto; não devemos afastar *a priori* a possibilidade de que tenham algum meio de acesso à acepção literal dessas formas. Por outro lado, não é impossível que interfira aí a instrução gramatical escolar, o que de certo modo viciaria o argumento. Entre tanto, na falta de coisa melhor, podemos admitir que há uma possibilidade limitada de se aferir intuitivamente o grau de contextualização — e, daí, o sentido literal — das formas lingüísticas.

Outra se pensar em um tratamento separado da semântica e da pragmática é a existência de traços de significado que independem do contexto extralingüístico para sua depreensão segura. Por exemplo, as afirmações abaixo são válidas em virtude exclusivamente das formas lingüísticas mencionadas, sem consideração da situação:

(a) O objeto direto nunca se refere ao agente de uma ação.

(b) O adjetivo modifica o significado do substantivo que se constrói com ele no mesmo sintagma.

(c) Os sintagmas nominais, e só os sintagmas nomi nais, podem ser utilizados para denotar entidades específicas do mundo real.

(d) Um sintagma nominal que seja sujeito de um verbo no futuro do pretérito pode ser interpretado como não se referindo a uma entidade real, mas apenas a uma entidade hipotética cuja existência não é pressuposta. Assim, em:

(24) uma pitanga doce me surpreenderia não há pressuposição nenhuma de que pitangas doces realmente existam. Mas, se o verbo estiver no pretérito perfeito, esse tipo de interpretação torna-se impossível:

(25) uma pitanga doce me surpreendeu

Esta frase engloba necessariamente o pressuposto de que pitangas doces existem no mundo real.

Como essas, há muitas outras afirmações relativas ao significado que não dependem do contexto. Evidentemente, esse fato não é suficiente em si para autorizar uma sepa ração teórica entre semântica e pragmática. Essa separação depende de ser ou não possível distinguir um nível em que as sentenças completas possam ser interpretadas sem referência ao contexto; e nos exemplos acima só lidamos com traços isolados de significado. Mas, pelo menos, já não se pode dizer que não seja possível em absoluto atacar o problema do significado com referência exclusiva às formas lingUísticas.

A questão não fica resolvida com isso; nem há, que eu saiba, argumentos conclusivos em favor de uma das opiniões em conflito. Como já dei a entender anteriormente, a separação entre pragmática e semântica simplifica o trabalho de elaboração da gramática, pois dispensa a con sideração do contexto extralingüístico (com toda a sua complexidade, ainda não capturada em nenhuma teoria) na descrição do significado. Isso significa, claro, que parte do significado final das formas lingüísticas ficará sem ser descrita, ou só será descrita informal e incompletamente — preço que pagaremos, já que é forçoso assumir uma posição, e não há à vista alternativa mais atraente. Teremos, pelo menos, a vantagem, sobre os estudos tradicionais, de estarmos cônscios da fragilidade dos fundamentos da nossa posição.

A descrição do significado na gramática

Partindo, pois, da definição de semântica como o estudo do significado das expressões lingüísticas maximamente descontextualizadas, e da pragmática como o estudo do seu significado em contexto, a primeira

pergunta a ser feita é: Até que ponto poderemos prescindir de informação pragmática na redação da gramática?

A resposta dependerá da seleção que fizermos dos temas a serem descritos. Seria possível selecionar um conjunto de temas tal que mesmo a semântica *stricto sensu* acabasse sendo excluída da gramática. Mas isso não seria indicado, por mais de uma razão. Primeiro, seria desejável cobrir, na medida do possível, o mesmo conjunto de assuntos tradicionalmente abordado pelas gramáticas tradicionais; essa decisão é, quando nada, boa política, pois facilita a comparação entre a GT e a nova proposta. Depois, o estudo do significado é em si suficientemente importante e interessante para merecer atenção, ainda que à custa de incertezas e falta de precisão. Assim, sou de opinião que se deve incluir, de alguma forma, o estudo do significado em uma gramática pedagógica.

Antes, porém, de tratar da seleção dos tópicos a serem incluídos (o que não é objetivo do presente texto), pode ser interessante fazer um levantamento sumário dos tipos de informação referentes ao significado que aparecem usualmente nas gramáticas. Com esse inventário, apuraremos até que ponto problemas de ordem pragmática são tradicionalmente incluídos nas gramáticas.

Examinando algumas gramáticas, verifica-se que as afirmações relativas ao significado são de muitos tipos diferentes. Assim, encontramos os seguintes tipos de afirmações:

(a) *Denotação*

(Isto é, afirmações acerca do sentido ou da referência de elementos gramaticais ou léxicos.) Por exemplo:

"Substantivos nomeiam seres."

"O sujeito indeterminado não se refere a uma pessoa determinada, por se desconhecer quem executa a ação ou por não se querer nomeá-lo."

"Verbos denotam ações."

Esse tipo de afirmação, por descrever fenômenos in dependentes do contexto, é geralmente incluído na semântica.

(b*) Funções temátic*as (segundo a terminologia de Gruber, 1965, e Jackendoff, 1972).

(Trata-se de afirmações referentes às relações semânticas entre os diversos constituintes da sentença ou do sintagma.) Exemplos:

"O sujeito pode ser agente, paciente. .

"O adjetivo qualifica o substantivo."

Aqui, novamente, temos afirmações independentes do contexto, pelo que se incluem na semântica.

(c) *Dêixis*

Exemplos:

"Este indica aquele que está perto de quem fala."

“Os pronomes se referem às pessoas do discurso.”

“O presente do indicativo enuncia um fato que ocorre no momento em que se fala.”

Aqui já temos que considerar fenômenos pragmáticos, segundo nossa definição. Com efeito, o significado (mais precisamente, a referência) de um elemento dêitico não pode ser depreendido ou estudado independentemente do contexto em que ocorre. *Eu* se refere a Maria ou a João, segundo quem fale no momento.

(d) *Força ilocucionária*

(Ou seja, afirmações que distinguem perguntas, declarações, ordens, exclamações etc.) Exemplos:

“O vocativo serve para invocar, chamar.”

“As interrogativas se usam para expressar falta de informação sobre um assunto específico, e para pedir ao ouvinte que forneça verbalmente essa informação.”

Nada obstaria, em princípio, a que a força ilocucionária fosse considerada um traço semântico das formas lingüísticas, associado a essas formas sem dependência do contexto. Por exemplo, se a segunda afirmação acima fosse verdadeira, isto é, se toda e qualquer sentença interroga tiva (terminada em “?”) tivesse a força ilocucionária de “pedido de informação verbal”, poderíamos dizer que esse é um traço de significado inerente àquela estrutura, independentemente do contexto em que possa ocorrer. O caso da força ilocucionária é, como se vê, diferente do da dêixis, que, por sua própria natureza, não pode ser interpretada sem o contexto.

No entanto, observa-se que a força ilocucionária é mais freqüentemente depreendida a partir da forma lingüística mais o contexto do que a partir da forma lingüística pura. Em outras palavras, na maioria dos casos o papel do contexto é crucial para que se perceba a força ilocucionária de um enunciado.

Assim, uma frase como:

(26) você gosta de caramelo?

será interpretada como um pedido de informação sobre os gostos do interlocutor, ou então como uma oferta de caramelos, conforme o contexto em que apareça.

Assim, para nossos objetivos, é melhor estabelecer que, embora a força ilocucionária seja talvez um traço basicamente semântico, não é possível estudá-lo na prática sem considerações de ordem pragmática.

(e) *Tratamento*

(É o caso de formas cujo uso adequado depende do reconhecimento de relações sociais entre os interlocutores.)

Exemplo:

“*Você* é forma de intimidade; o *senhor* é forma de respeito e cortesia.”

A referência desses pronomes, propriamente dita, é sem dúvida um problema de pragmática (dêixis): podemos repetir o que dissemos acima sobre o pronome *eu*: você será Maria ou João, dependendo de quem é o ouvinte. Mas que dizer do fato de que *você* seja familiar, o *senhor*, formal? Essa informação está codificada nas próprias formas lingüísticas em questão; creio que se pode considerar um tipo de denotação, a ser classificado como fenômeno semântico.[2]

(f) *Status funcional*

(Trataremos desses casos, com vagar, mais adiante; por ora, basta notar que um dos *status* funcionais é o de tópico de uma sentença, ou seja, o elemento sobre o qual se faz uma declaração.) Exemplo:

"Sujeito é o termo sobre o qual se faz uma declaração."

Creio que aqui não temos nem estritamente semântica nem estritamente pragmática. Trata-se do componente a que Chafe (1976) chamou "embalagem" (*packaging*), e os lingüistas da escola de Praga, *perspectiva funcional.*

Até onde pude verificar, essas seis categorias cobrem todas as afirmações referentes ao significado encontradas na GT. Como se vê, é um grupo bastante heterogêneo, mas as afirmações semânticas predominam. Se nos limitarmos à seleção tradicional de tópicos, poderemos restringir a consideração do contexto ao caso da dêixis (onde o papel do contexto é relativamente claro e delimitado) e ao da força ilocucionária que, ainda assim, parece admitir um tratamento parcial em bases semânticas. Quanto ao *status* funcional, creio que vale a pena incluí-lo na descrição; para maiores detalhes, ver a subseção seguinte.

Que se pode concluir dessas observações? Pelo menos alguns pontos estão razoavelmente claros. A descrição dos aspectos ligados ao significado a ser incluída na gramática poderá ser discriminada em três categorias: (a) os aspectos *semânticos stricto sensu* serão descritos em termos de propriedades semânticas de estruturas lingüísticas. (b) Os aspectos *pragmáticos* serão descritos em termos de interpretações semânticas, mais traços do contexto situacional, e se restringirão a certas áreas, notadamente a dêixis e a força ilocucionária. (c) Finalmente, os aspectos ligados ao *status* funcional serão descritos em termos de análise do discurso. O resultado final, é claro, será um tanto heterogêneo em grau de precisão, exaustividade e mesmo clareza. Não ve necessariamente um problema: uma gramática moderna não pode deixar de refletir o estado da lingüística moderna.

Traços discursivos na descrição gramatical

Vamos examinar nesta subseção o problema da inclusão de traços da análise do discurso na descrição gramatical (entendendo-se por "análise do

---

discurso” qualquer estudo da organização da expressão lingüística além do nível do período).

A GT geralmente não pretende considerar fenômenos ligados à organização do discurso, estabelecendo o período como domínio máximo da análise. Essa é sem dúvida uma opção prudente, dado o estado incipiente dos estudos da organização do discurso. Embora haja trabalhos de importância nessa área, o progresso tem sido difícil porque, ao contrário do que se dá com a sintaxe, ainda não se chegou a uma teoria coerente que possa guiar a investigação em análise do discurso. Desse modo, os resultados, ainda que muitas vezes interessantes, não chegam a estruturar-se em uma visão geral do fenômeno.

Não obstante, a gramática talvez não possa escapar totalmente de levar em conta certos aspectos da análise do discurso, importantes para a descrição de alguns processos tradicionalmente considerados gramaticais, tais como: o emprego do artigo definido frente ao indefinido; as condições de uso dos pronomes anafóricos; ou a determinação do “elemento sobre o qual se faz uma declaração”. Ao considerar esses pontos, a GT se afasta de sua decisão tácita de limitar-se ao período, fazendo incursões breves e informais no domínio do discurso.

Os pontos mencionados acima se relacionam com o chamado “aspecto funcional” da linguagem; começarei, portanto, por conceituar brevemente essa noção.

Assim como a gramática se ocupa (na sintaxe) dos *status* formais dos sintagmas, tais como o de sujeito ou o de objeto direto, ela também deverá ocupar-se (em outro componente, separado da sintaxe e talvez da semântica) dos *status* funcionais como o de “tópico” ou o de “foco de contraste”, que dizem respeito não propriamente à forma ou ao conteúdo da mensagem, mas à maneira pela qual o conteúdo é apresentado no discurso. Por exemplo, podemos ter duas sentenças com a mesma análise sintática (no que diz respeito às funções dos termos) e a mesma interpretação semântica *stricto sensu*, mas que ainda assim apresentam diferenças que podemos levar em conta em uma gramática:

(27) chove muito em Belo Horizonte

(28) em Belo Horizonte chove muito

A função sintática e a interpretação semântica de *em Belo Horizonte* são as mesmas em (27) e em (28). Mas certamente a posição desse sintagma nas duas sentenças, além de merecer atenção por si mesma, correlaciona-se com uma diferença de forma de apresentação da informação. Dizemos que *em Belo Horizonte*, na sentença (28), tem o *status* funcional de “tópico”; o restante da sentença se denomina habitualmente “comentário” .

Que significa exatamente dizer que determinado ele mento é o tópico de uma sentença? Tradicionalmente se empregam termos como “ênfase”, “relevo”, “realce”, que requerem, eles próprios, uma explicação. Sem ter a pre tensão de chegar já a uma conceituação satisfatória, vou deter-me um pouco no problema da definição de “tópico”.

Examinamos, na seção 1, uma definição tradicional de sujeito que, como vimos, é inadequada:

(1) "O sujeito é o termo sobre o qual se faz uma declaração."

Embora a definição (1) não seja satisfatória para caracterizar -o habitualmente considerado na prática da analise, ela capta um aspecto real da interpretação das sentenças. É fato, em muitos casos, que a sentença veicula informação primariamente sobre um de seus termos. Assim, creio que se pode dizer que a diferença entre (27) e (28) é que em (28) a informação se apresenta basicamente sob a forma de uma afirmação acerca de Belo Horizonte, ao passo que (27) é mais neutra quanto a isso. Vejam-se tambem os exemplos seguintes

(29) Carmelita beliscou o paroco

(30) o paroco, Carmelita beliscou

(31) o sapato não está debaixo da cama

(32) debaixo da cama, o sapato não esta

Tarnbém se encontram as designações tema (="tópico") e rema (="comentário")

Nesses exemplos, a maioria das pessoas concordaria que as sentenças (30) e (32) apresentam informação pri mariamente "sobre" o pároco e "sobre" debaixo da cama; por outro lado, é claro que também há informação sobre Carmelita em (30) e sobre o sapato em (32), mas isso parece ficar um pouco na sombra. Em (30), por exemplo, é como se o falante avisasse, através da colocação de o pároco no início, que pretende falar do pároco mais do que de qualquer outra personagem.

Já nos exemplos (29) e (31) não me parece que haja um elemento tão nitidamente privilegiado; assim, (29) não é tão claramente uma afirmação "sobre" Carmelita quanto (30) é uma afirmação "sobre" o pároco. Othon Garcia parece ter intuição semelhante, quando diz

"Se, pela ordem direta, o objeto direto, o objeto indireto e o predicativo se pospõem ao verbo, basta antepô-los para que eles, por ocuparem uma posição insólita, ganhem maior relevo." (GARCIA, 1975, p. 248.)

Se chamarmos de "tópicos" os sintagmas o *pároco* em

(30) e *debaixo da cama* em (32), poderemos modificar (1) como se segue:

(33) O tópico é o termo sobre o qual se faz uma declaração.

Essa pode ser considerada uma primeira aproximação à definição de "tópico" em termos funcionais, isto é, em termos da contribuição que ele dá ao andamento do discurso.

Além da definição funcional, será preciso também procurar uma definição formal de "tópico": uma enumeração dos recursos sintáticos de que dispõe a língua para marcar determinado termo como tópico. Vimos, pelos exemplos dados, que um desses recursos é o de colocar o termo topicalizado no início da frase; isso pode ser conseguido através de uma simples

transposição de ordem, como nos exemplos vistos, ou através de outros processos da língua, como a passivização. Por outro lado, sabemos que o tópico é mais nítido em (30) e (32) do que em (29) e (31). Assim, se quisermos caracterizar formalmente o tópico pela sua posição no início da sentença, teremos de admitir que seu valor funcional varia em força, sendo mais forte no caso do objeto direto do que no do sujeito, por exemplo. A passagem citada, de Garcia (1975), parece sugerir, antes, uma caracterização não em termos da posição sintática apenas, mas também em termos de desvio a partir de uma posição "normal", que corresponderia à chamada "ordem direta". Esses problemas são interessantes, mas não será possível abordá-los no presente texto.

Voltemos à definição funcional de "tópico". Em (33) temos, como vimos, uma primeira aproximação, ainda bastante vaga. A seguinte passagem deve contribuir para deixar a questão um pouco mais clara:

"O que os tópicos parecem fazer é limitar a aplicabilidade da predicação principal a um certo domínio restrito. (...) o tópicb'—estahelece uma moldura espacial, temporal ou individual dentro da qual a predicação principal é válida."

(CHAFE, 1976, p. 50.)

Um ingrediente que aparece com freqüência associado ao tópico é a contrastividade. Assim, em

(32) debaixo da cama, o sapato não está

entende-se, em geral, que *debaixo da cama* se opõe a outros lugares onde o sapato possa estar. No entanto, o tópico não é sempre necessariamente um foco de contraste. Por exemplo, *em Belo Horizonte*, em

(28) em Belo Horizonte chove muito

ainda que seja claramente tópico, não é contrastivo. Desse modo, os status de "tópico" e "foco de contraste" devem ser mantidos distintos.

A literatura lingüística distingue um bom número de *status* funcionais, e nem sempre oferece definições rigorosas que permitam a sua manipulação segura. Além da oposição tópico/comentário que acabamos de examinar, têm importância especial os três *status* seguintes:

(a) *foco de contraste*: o elemento mencionado dentre vários possíveis candidatos que podem estar na mente do interlocutor:

(34) foi o cachorro que sujou o tapete e não, digamos, o canário ou o gato.

(b*) Definido (versus indefinido)*: *status* atribuído aos elementos que o falante supõe que possam ser identificados sem ambigüidade pelo ouvinte. Esse *status* é um dos fatores presentes na determinação do uso do artigo definido em oposição ao indefinido.

(c*) Dado (versus novo)*: elemento que o falante presume que esteja ativado, no momento, na memória do interlocutor (e não simplesmente *conhecido* pelo interlocutor). Somente elementos *dados* poderiam ser retomados através de pronomes pessoais. Assim, uma frase como:

(35) ele acabou de me telefonar

só é compreensível se o ouvinte estiver com o referente de *ele* ativado, presente em uma espécie de "linha de frente" na sua memória; por exemplo, se (35) for dita imediatamente depois de:

(36) você se lembra do Ronaldão?

Esses diversos *status* não são mutuamente exclusivos; devem ser encarados, antes, como traços que em princípio podem superpor-se livremente, de tal maneira que um sintagma possa ser ao mesmo tempo tópico, foco de contraste, dado e definido, por exemplo. Há, é verdade, certas restrições às possibilidades de associação dos diversos *status*, mas não são suficientemente conhecidas no atual estágio da investigação.

Para uma discussão de todos esses *status* e de várias questões relacionadas, no contexto do português, recomendo a leitura de Liberato (1980); certos aspectos da dicotomia dado/novo e sua representação formal são desenvolvidos por Fulgêncio (1983). Evidentemente, há ainda muito o que fazer antes de atingirmos uma concepção toleravelmente clara dos *status* funcionais e de seu papel no andamento do discurso. É inegável, contudo, que se trata de um aspecto importante da linguagem, certamente merecedor da atenção do estudioso.

A elipse e os constituintes vazios

Deixemos agora a discussão dos fenômenos semânticos e funcionais para nos voltarmos para a questão da descrição sintática. Um dos primeiros problemas que se apresentam quando se procura planejar uma descrição coerente da sintaxe da língua é o de se se deve ou não admitir a postulação de elementos elípticos ("subentendidos") válidos para a análise gramatical.

Consideremos o período seguinte:

(37) Serafim toca flautim, e Ivone trombone

Serafim toca flautim é, evidentemente, uma oração; mas como classificar *Ivone trombone*? Qual é a função de *Ivone*? E a de *trombone*? Uma solução tradicional é a de "subentender" um verbo entre *Ivone* e *trombone*: uma cópia do verbo da primeira oração, *toca*. Desse modo, a seqüência *Ivone trombone* se analisa como uma oração, *Ivone* como sujeito e *trombone* como objeto direto. Essa posição tem vantagens e também perigos, conforme veremos a seguir.

Há vantagens evidentes em se considerar a seqüência *Ivone trombone* uma oração (ainda que "reduzida"). Vere mos essas vantagens mais adiante, mas algumas já podem ser mencionadas: (a) Essa seqüência se encontra, em (37), coordenada por meio de e a uma oração; ora, sabemos que e sempre coordena elementos (itens léxicos, sintagmas) pertencentes à mesma classe. (b) Essa seqüência recebe uma interpretação semântica típica de oração (paralela à de *Serafim toca flautim*). (c) Além disso, se não for uma oração, o que é que vai ser?

Mas para classificarmos essa seqüência como uma oração também teremos de enfrentar certos problemas. Se não quisermos introduzir um elemento subentendido, será preciso admitir que uma oração pode ser

composta de dois sintagmas nominais justapostos, o que vai complicar a gramática. Basta dizer que será preciso estabelecer em que circunstâncias uma seqüência de dois sintagmas nominais pode ser uma oração, o que não será nada fácil.

Introduzir elementos subentendidos válidos para a análise, como o verbo *toca* na segunda oração de (37), só se pode fazer tomando precauções contra certas possibilidades perigosas. É preciso evitar a introdução desenfreada de elementos, ainda que estes sejam inegavelmente subentendidos. Para tomar um caso concreto, não podemos introduzir *Zé morreu* na análise sintática da sentença:

(38) o jegue matou Zé

Não há dúvida de que (38) deixa entendido que Zé morreu. No entanto, levar em conta essa informação para uma análise *sintática* (isto é, formal) nos levará ao caos.

A saída adotada por Jackendoff (1972) e pelos gerativistas atuais é permitir essa introdução em circunstâncias cuidadosamente controladas. Vamos postular que uma seqüência como (37) apresenta uma lacuna (representada aqui por "Ø") no lugar que corresponderia ao verbo da segunda oração:

(39) Serafim toca flautim, e Ivone 0 trombone

Em algum lugar da gramática haverá uma regra (ou um princípio mais geral) estabelecendo algum tipo de identidade semântica entre a lacuna e um elemento presente no restante do período e estritamente definido pela mesma regra. No caso de (37), a relação se faz entre a lacuna e o verbo da primeira oração, *toca*. Assim, restitui-se o significado 'toca' na segunda metade de (37), ficando expresso na gramática o fato de que *Ivone trombone* se interpreta de maneira paralela à de *Serafim toca flautim*, em (37).

A questão se coloca, então, da seguinte forma: Em que circunstâncias será permitida a introdução de elementos vazios tais como o 0 de (39)?

Observe-se que a introdução de um elemento vazio não pode ser contemplada apenas a nível do significado; será preciso a presença de tais elementos a nível da sintaxe. Se só admitirmos a restituição de um significado (toca), sem a introdução de um nódulo sintático, nem todos os problemas ficarão resolvidos. Continuaremos, por exemplo, a ter a estrutura sintática composta de dois sintagmas nominais seguidos, comportando-se como uma oração. A saída é postular, a nível sintático, a presença de nódulos vazios, isto é, substantivos, verbos etc., *não*

Não será possível aqui discutir que tipo, exatamente, de identidade semântica é necessário. Não se trata de correferência pura e simples, pelo menos por causa de frases como:

mandei vir uma *Correinha* e ela também pediu uma 0 onde claramente se trata de duas *Correinhas*, não uma só; mas aí também necessitamos de uma regra que relacione a lacuna final com *Correrinha* da primeira oração.

preenchidos lexicamente. Assim, (37) acabaria sendo re presentada, ainda a nível da sintaxe, da seguinte maneira :

(40) Serafim toca flautim, e Ivone trombone

O nódulo vazio em (40) é um verbo, conforme indica o rótulo "V"; é um verbo tal como toca, com a diferença de que não está preenchido lexicamente; não tem realização fonológica, em suma. Assim, as duas metades de (40) são sintaticamente paralelas, o que explica o seu paralelismo semântico e seu comportamento sintático.

Ao adotarmos em nossa análise esse recurso, estamos, pela primeira vez, lançando mão de entidades que se po dem classificar de "abstratas". Isto é, estamos admitindo a existência de constituintes válidos para a análise, mas que não têm realização fonológica; vale dizer, que não são diretamente observáveis. Antes de passar adiante, é importante justificar o uso de tais recursos. Só utilizaremos entidades abstratas se isso nos trouxer vantagens bem claras para a descrição gramatical; e, não é demais repetir, será também necessário encontrar critérios de controle estrito da introdução de tais elementos abstratos, para evi tar situações análogas à que vimos em conexão com o exemplo (38).

Vejamos primeiro a questão de justificar o emprego de elementos vazios. Que vantagens nos traz a adoção de tais análises abstratas?

Observarei, antes, que uma análise como a ilustrada em (40) só é novidade à primeira vista. A análise tradi cional afirma que *trombone*, em (37), é um objeto direto; e que não é o objeto direto do verbo da primeira oração (isto é, *toca* da primeira oração só tem um objeto, *flautim*). Logo, deve ser objeto direto de algum outro verbo; não

O 'V', 'N' etc. subscritos são usados para indicar a classe a que pertence o nódulo.

Havendo outro verbo explícito em (37), forçosamente te mos de concluir que a análise tradicional também postula a presença de um "verbo" abstrato na segunda oração. A diferença está em que isso não é suficientemente explicitado na GT.

Agora passemos às vantagens que nos traz a análise ilustrada em (40). Algumas já foram mencionadas, e começarei por recapitulá-las:

(a) *Interpretação semântica*: a presença de um verbo (lexicamente vazio) na segunda metade de (40) pode servir de base para a operação das regras de interpretação semântica, que deverão dar à segunda metade da seqüência uma acepção paralela à da primeira metade. Como já apontei, aqui deveremos ter uma regra (ou princípio geral) que relacione o nódulo vazio com o constituinte correspondente na primeira metade da seqüência, interpretando ambos como semanticamente idênticos. Ora, essa regra ou princípio, ao introduzir o verbo vazio na segunda metade de (40), torna possível o processamento paralelo das duas metades (qu estruturalmente semelhantes). Já se não tivéssemos o verbo vazio, seria preciso formular uma regra diferente para interpretar a segunda metade, de modo que, apesar de seus resultados semanticamente semelhantes, as duas metades seriam interpretadas por mecanis mos distintos. Como se vê, a presença do elemento abstra to possibilita a formulação de um componente semântico

mais consistente e econômico (e também, parece-me, mais satisfatório intuitivamente).

(b) *Coordenação a estruturas paralelas*: a análise de (40) nos permite classificar a segunda metade da estrutura como "oração", tal como a primeira. Isso faz com que o fato de elas estarem coordenadas através de *e* deixe de ser estranho: aqui, como em geral acontece, a conjunção e está coordenando itens pertencentes à mesma classe.

(c) *Enquadramento em uma estrutura típica da língua*: com a análise de (40) já não é preciso especificar a existência de uma estrutura composta de dois sintagmas nominais seguidos e que não é, ela mesma, um sintagma nominal. Agora, *Ivone trombone* é a manifestação super ficial de uma oração, estruturada como toda e qualquer oração, com a diferença única de que contém um verbo vazio.

Como:

Além desses argumentos, podem ser aduzidos outros,

(d) *Presença de sujeitos e objetos:* acontece que uma seqüência como *Ivone trombone* em (37) funciona como uma oração de mais de uma maneira. Por exemplo, o primeiro sintagma nominal tem as características de um sujeito, e o segundo as características de um objeto. Para ver isso, tome-se a frase:

(41) Totó beijou Lassie, e Maria Pedrinho

Se substituirmos o nome *Maria* por um pronome pessoal, veremos que a forma reta (nominativa) é exigida:

(42) Totó beijou Lassie, e eu Pedrinho

Mas se fizermos o mesmo com *Pedrinho*, será preciso colocar uma forma oblíqua (objetiva):

(43) Totó beijou Lassie, e Maria a mim

Esses fatos decorrem automaticamente da teoria de que os dois sintagmas nominais são o sujeito e o objeto de uma oração cujo verbo é lexicamente vazio. Já se não postulássemos a existência desse verbo, teríamos de des crever e explicar separadamente o comportamento dos pro nomes em (42) e (43).

6 Aqui não pode ocorrer a forma me, que só aparece ao lado de um verbo explícito (por ser um clítico).

(e*) Reposição de itens elípticos*: pelo menos em certos casos, é possível recolocar no lugar do nódulo vazio uma cópia do seu antecedente, sem que se observe mudança sensível na interpretação semântica ou no compor tamento sintático da seqüência. Assim, ao lado de (37), podemos ter:

(44) Serafim toca flautim, e Ivone toca trombone

Esse fato seria misterioso se *Ivone trombone* fosse analisado corno uma estrutura independente, não-oracional; mas decorre naturalmente da análise dessa seqüência como uma oração.

É possível encontrar mais outros argumentos em favor do tipo de análise que estamos examinando; mas os cinco argumentos dados acima deverão ser suficientes para nos convencer da conveniência da introdução de nódulos vazios em certos casos.

Nos exemplos acima, os casos vistos são todos para lelos a (37), onde se suprime o verbo de uma de duas orações coordenarlas. Mas a análise deverá generalizar-se para outros casos; talvez precisemos de substantivos vazios:

(45) comprei uma vaca amarela e uma 0 malhada

(46) Juca dorme com a dentadura e Manuel dorme sem 0

(Neste último exemplo, é interessante introduzir um nome

— ou um sintagma nominal, aqui não importa muito — logo depois da preposição *sem*; isso nos ajudará a simplificar a gramática, porque evitaremos ter de admitir casos excepcionais de preposições não seguidas de nome. Pode remos dizer que preposições são sempre seguidas de sin tagma nominal, oracional ou não, sem exceção.)

É claro que as análises poderão ser muito diferentes de caso para caso, pois as lacunas estão longe de terem as mesmas propriedades. Mas terão em comum o fato de que algum tipo de nódulo vazio se faz necessário.

Esses exemplos foram dados para ilustrar como se pode utilizar a análise com nódulos vazios a fim de des crever com propriedade certos fatos da língua. Creio que a argumentação é suficiente para justificar a decisão de incluir esse tipo de recurso na descrição gramatical. Agora passemos ao problema do controle da introdução de nódu los vazios.

A necessidade de postular a presença de nódulos vazios (constituintes não-realizados fonologicamente), em favor da qual argumentei acima, coloca o problema de decidir quando é lícito recorrer a eles. Pelas razões que veremos, as condições de introdução de nódulos vazios precisam ser estritamente controladas.

Suponhamos, para começar, que não haja controle nenhum, de modo que possamos introduzir elementos irrestritamente. Nesse caso, será permitido colocar nódulos va zios de *qualquer* classe em *qualquer* posição na oração. Ora, isso tornaria impossível uma análise coerente, pois a teoria possibilitaria a construção de um número ilimitado de alternativas a qualquer análise que viesse a ser proposta. Dessa maneira nunca seria possível decidir se uma análise é melhor do que outra, já que todas estariam de acordo com a teoria, não havendo verdadeiro compromisso com os fatos da língua.

Mas, evidentemente, ninguém nunca defendeu a introdução irrestrita de nódulos vazios. Explícita ou implicitamente, sempre se colocam restrições. Por exemplo, poderíamos estipular que o material introduzido deveria obriga toriamente ter correlação com a interpretação semântica da sentença; ou seja, só introduziríamos material "subentendido", à maneira tradicional; e a única restrição seria essa. Em muitas passagens, a GT parece seguir esse princípio, que já pode ser considerado relativamente

severo, pois impõe restrições bem definidas à postulação de elementos vazios; mas ainda assim o princípio é insuficientemente severo, conforme veremos. Ele permite a introdução de um nódulo vazio, classificado como “verbo”, na segunda metade de (37), de modo que temos:

(40) Serafim toca flautim, e Ivone Ø trombone

A introdução do verbo vazio se justificaria simples mente pelo fato de que a interpretação semântica da sen tença contém um segundo “toca”, não representado por nenhum item explícito.

No entanto, se nos limitarmos à restrição de só introduzir material semanticamente subentendido, poderemos também introduzir Zé *morreu* como parte da estrutura de:

(38) o jegue matou Zé

Seria, então, o caso de analisar Zé como sujeito (de *morreu*), o que pode talvez fazer sentido em termos da interpretação semântica, mas certamente seria inaceitável dentro de urna sintaxe compreensível da língua. Não obstante, é inegável que ‘Zé morreu’ é parte da interpretação semântica de (38).

Para tomar outro exemplo (dentre uma multidão), o princípio da introdução de elementos vazios sempre que isso se justificar semanticamente impede a distinção entre predicativos e adjuntos adnominais. Isso porque um sintagma como:

(47) a mula manca

onde manca é um adjunto adnominal, poderia ser analisado como sendo:

(48) a mula que é manca

onde m*anca* é predicativo, pois (47) e (48) são semanticamente semelhantes.

A GT parece, às vezes, partir desse princípio. Isso não se encontra em todos os autores, nem sistematicamente em nenhum deles; mas está subjacente a análises como a seguinte. Muitas vezes se argumenta que o sujeito de:

(49) vende-se este pardieiro

é *este pardieiro*, porque (49) “corresponde” a:

(50) este pardieiro é vendido

onde *este pardieiro* é sujeito. Sem discutir qual é o sujeito de (49), nem a questionável sinonímia entre (49) e (50), vejamos a estrutura do argumento. Primeiro, não é muito claro o que significa dizer que uma estrutura “corresponde” a outra; mas suponhamos que se trate de uma seme lhança semântica de alguma espécie, como aquela que utilizamos acima para restringir a introdução de elementos vazios. Assim, uma estrutura “corresponde” a outra quando elas são sinônimas (em alguma medida, a ser explicitada). Nesse caso, o argumento subjacente à análise de (49) segue o seguinte raciocínio: quando uma sentença A tem uma interpretação

semântica suficientemente semelhante à de outra sentença *B*, então a análise sintática de B vale também para *A*. Logo, se (49) é semanticamente semelhante a (50), o fato de *este pardieiro* ser o sujeito de (50) é razão suficiente para se concluir que *este pardieiro* é também o sujeito de (49).

Se quisermos aplicar estritamente esse critério, acaba remos indo muito mais longe do que desejamos. Só se explica a sua presença (subjacente) na análise tradicional pela informalidade com que os recursos de análise são utilizados, sem que haja muita consciência de suas implicações. O fato é que, se minha interpretação está correta, a análise de (49) com base em (50) pressupõe princípios que nos permitem analisar este pardieiro como o sujeito de (51):

(51) estão vendendo este pardieiro

ou, conversamente, analisar este pardieiro como objeto di reto de (49) ou mesmo de (50), já que ambas "correspondem" a (51). É fácil ver que esse caminho não nos levará a uma análise coerente.

Já que a restrição (semântica) vista acima é insufi cientemente estrita, permanece o problema de delimitar os casos em que se pode permitir a introdução de nódulos vazios válidos para a análise sintática. Como chegar a uma restrição formulada de tal modo que favoreça uma análise sintática tão simples e sistemática quanto possível; que possibilite regras de interpretação semântica gerais e não *ad hoc*; e que seja, a própria restrição, simples, coerente e bem motivada? Essa é uma questão muito atual em Lingüística, o que significa que não há respostas verdadeiramente satisfatórias a ponto de serem aceitas pela maioria dos pesquisadores. Aqui não será possível chegar a uma formu1ação sequer aproximada, dessa restrição (provavelmente um princípio, ou um conjunto de princípios gerais da organização da língua). Teremos de nos conten tar, por ora, com o reconhecimento da natureza do problema; consolemo-nos com o pensamento de que localizar um problema já é um grande passo para sua solução.[3]

A explicitação detalhada das condições que devem governar a introdução de nódulos vazios tem sido objeto da atenção intensiva de muitos lingüistas. Apesar de não haver, como apontei acima, uma resposta universalmente aceita, já existem algumas propostas muito específicas. Uma das primeiras tarefas a serem empreendidas quando da elaboração da nova gramática será a seleção criteriosa de uma dessas propostas e sua adaptação às necessidades imediatas da descrição do português.

Classes de palavras

Concentremo-nos agora no problema da classificação das palavras — mas tendo em mente que a noção de classificação se estende a unidades maiores (e menores) do que a palavra.

Antes de enfrentar alguns dos muitos problemas liga dos à classificação das palavras, é interessante discutir previamente a conveniência de se considerar a palavra como unidade básica da análise, inclusive porque tal conveniência tem sido contestada por alguns lingüistas.

Martinet, seguindo um uso freqüente entre os estruturalistas, não concede à palavra nenhum *status* especial na gramática, preferindo ater-se com exclusividade aos morfemas (que ele denomina *monèmes*). Diz ele:

"não se trata de limitar-se às 'palavras', isto é, aos segmen tos do enunciado facilmente isoláveis, tratando em bloco aquilo que é, formalmente, difícil de analisar. Não há ne nhuma razão válida para tratar diferentemente as relações do 'artigo definido' com o 'substantivo' e as do 'futuro' com o 'verbo', com o pretexto de que 'artigo' e 'substantivo' são separáveis (...) enquanto que 'futuro' e 'verbo' não o são" (MARTINET, 1 p. 10.)

Martinet considera a palavra uma unidade clandestina, estabelecida sobre bases arbitrárias e teoricamente injustificáveis. Neste trabalho preferirei adotar posição oposta, admitindo (juntamente com a GT) a necessidade de se distinguir a palavra como unidade de análise; com isso se estabelece também uma distinção nítida entre sintaxe e morfologia.

Como justificar essa posição? Em fonologia parece inevitável a aceitação da palavra como unidade. Lembrarei apenas fenômenos bastante bem conhecidos como o posicionamento do acento tônico, a redução e ensurdecimento das vogais e a ocorrência de vogais tônicas orais diante de consoante nasal (compare-se a pronúncia de cano torto com a de cá no Torto), todos os quais dependem dos limites de palavra para serem descritos. Assim, ainda que se pudesse prescindir da unidade "palavra" em sintaxe, ela ainda seria crucial dentro da gramática, pela necessidade de descrever tais fatos fonológicos.

No trecho citado acima, Martinet insinua que a postulação da palavra como unidade vem do fato de que a análise da distribuição dos morfemas dentro da palavra é especialmente difícil (". . . tratando em bloco aquilo que é, formalmente, difícil de analisar"). A meu ver, já existe aqui um embrião de argumento em favor da palavra como unidade de análise: não seria o caso de haver uma dife rença qualitativa entre os princípios que regem a distri buição dos elementos a nível da frase ("sintaxe") e os que regem a sua distribuição a nível de palavra ("morfologia")? E não seria justamente essa diferença a responsável pela maior dificuldade que encontramos ao analisar a palavra, frente à frase? Os morfemas de uma palavra não se con como as palavras de uma frase: há diferenças radicais no que se refere à sua liberdade de ocorrência e movimentação. Dentro de uma palavra não se encontra nada de análogo às várias posições possíveis de um advérbio, como nos exemplos:

(52) Carlos chegou hoje

(53) Carlos hoje chegou

(54) hoje Carlos chegou

Ao contrário, a posição dos morfemas dentro da pa lavra é estabelecida com grande rigidez; se podemos ter, digamos,

(55) re-constru-çõe-s

nenhuma inversão é permitida entre os diversos morfemas; e essa situação é típica dentro da palavra.

Outro argumento em favor de se distinguir os dois níveis (palavra e frase) é a grande ocorrência de irregu laridades que se verifica na morfologia da palavra. Refiro-me a fenômenos como:

(a) Os verbos se dividem em classes segundo a vogal que precede os sufixos modo-temporais: *am-a-r, tem-e-r, part-i-r.* Essa divisão não se correlaciona com nada mais no comportamento gramatical dessas palavras.

(b) Vento muda o *t* do radical somente diante de

-aval: *vendaval.*

(c) -*aval*, que funciona como sufixo em *vendaval,* tem uma distribuição extremamente restrita, não podendo ser generalizado para nenhuma outra palavra; assim, não há *chuvaval,* nem *nevaval* etc.

(d) Dentre as formas que se podem incluir no para digma de ir, algumas apresentam o radical *i-,* outras *fo-*, outras *va-*, outras *v-* (a esse fenômeno se denomina "supletivismo").

Ora, nada de comparável se encontra no campo da sintaxe propriamente dita, onde as irregularidades são menores e de outra natureza. Creio que qualquer pessoa que tenha estudado a estrutura de uma língua como o português concordará comigo neste particular. Desse modo, uma tentativa de estabelecer uma "sintaxe" que vá, sem níveis intermediários, da sentença ao morfema nos forçará a lançar mão de dois conjuntos de regras muito diferentes, conforme estejamos descrevendo os fenômenos tradicionalmente chamados "sintáticos" ou os chamados "morfológicos". É melhor, sem dúvida, reconhecer logo essa dis tinção, incorporando-a à teoria.

Passemos agora a considerar a classificação formal das palavras: seu agrupamento em classes (sem levar em

A análise morfológica é a de Pontes (1973).

8 Ver uma discussão interessante do assunto em Camara, 1964,

p. 87 e segs.

conta seus traços semânticos). Sem negar a importância de uma classificação semântica, insisto na necessidade de separar os dois tipos de classificação, pelo menos como hipótese inicial de trabalho — para que não se tome como postulado algo que precisa ser investigado empiricamente, a saber, o grau de correspondência entre o aspecto formal e o semântico das expressões lingüísticas. A classificação formal das palavras (ou dos morfemas, ou dos sintagmas em geral) deve ser feita com base no seu comportamento sintático.

Antes de mais nada, para que classificar as palavras? Sabemos que o objetivo da separação das palavras em classes é permitir a descrição econômica e coerente de seu comportamento gramatical. Assim, não se coloca em uma gramática a afirmação de que a palavra *gato*, individual mente considerada, pode ser núcleo de um sintagma nominal; que a palavra *trinco* também pode ser núcleo de um sintagma nominal; e assim por diante, até enumerarmos todas as palavras que possuem essa propriedade. Em vez disso, estabelecemos, de uma vez por todas, que *gato*, *trinco* etc. são "substantivos"; e na gramática dizemos sim plesmente que os substantivos podem ser núcleos de sintagmas nominais. Agrupamos essas palavras em uma classe porque elas *têm comportamento sintático semelhante*.

Ora, daí decorre que o estabelecimento das classes de palavras deve ser feito (dentro da sintaxe) com base no estudo do comportamento sintático das palavras, e somente nele: duas palavras serão incluídas na mesma classe se e somente se tiverem comportamento sintático semelhante. Por "comportamento sintático" entende-se o conjunto das relações que a palavra pode ter com as demais formas dentro da sentença. Por exemplo, são parte do comportamento sintático da palavra *casa* as seguintes propriedades:

(a) pode ocorrer como núcleo de um sintagma nominal .

(b) Pertence a um paradigma cujos membros se opõem quanto ao número.

(c) Não concorda em gênero com o núcleo de seu sintagma.

(d) Pode ocorrer logo após um artigo, formando a seqüência um sintagma nominal.

É esse tipo de propriedade que devemos investigar ao discutirmos a que classe atribuir a palavra *casa*. Outras propriedades dessa palavra, ainda que corretamente expressas, podem ser inteiramente inúteis para nossos obje tivos. Desse modo, não interessa saber que *casa* nomeia um objeto físico, em oposição a *tristeza*, porque *casa* e *tristeza* se comportam, sintaticamente, de maneira semelhante (evidentemente, as posições se invertem se nosso objetivo for descrever a semântica da língua).

Digamos que queremos investigar se se deve colocar na mesma classe as palavras *branco, casa, ele, corremos e sem*. Começaremos por selecionar um critério formal (uma propriedade sintática ou morfológica), que servirá para iniciar o trabalho. Pode ser o seguinte:

(56) A propriedade de poder ser núcleo de um sin tagma nominal.

(Uma propriedade assim expressa pode ser chamada um *traço*.)

Agora tomemos as palavras em questão e verifiquemos se elas podem ou não ser núcleos de sintagmas nominais. Não é difícil ver que branco, casa e ele podem, corremos e sem não podem:

(57) o branco combina com qualquer cor

9 questão de como determinar o núcleo de um sintagma não é trivial. Ver a nota seguinte.

(58) esta casa está para cair

(59) ele me telegrafou domingo

Essas frases são todas perfeitamente aceitáveis. Mas não é possível encontrar frases onde *corremos* ou s*em* sejam núcleos de sintagmas nominais. Concluímos que *branco, casa e ele* são marcados positivamente quanto à propriedade (ou traço) (56). Já *corremos* e *sem* são marca dos negativamente quanto a esse mesmo traço. Isso pode ser indicado graficamente da seguinte forma:

(60) traço branco casa ele corremos sem

(56) + + +

Aí temos apenas o início da investigação da classificação dessas palavras. O início, parece-me, é bastante simples; não faltam complicações quando se começa a con siderar casos concretos em detalhe. A seguir, vou apresentar brevemente algumas dessas complicações, para ilustrar como se pode discutir, dentro de uma perspectiva formal, a classificação das palavras.

A partir do momento em que se admite que a definição das classes de palavras precisa ser feita de acordo com seu comportamento sintático, surge o problema das classificações contraditórias, isto é, das palavras que fica riam em uma classe segundo um critério, mas em outra segundo outro critério. Um exemplo é o dos adjetivos frente aos substantivos: deveremos colocar *branco* e *casa* na mesma classe, por poderem ambos ocorrer após um artigo, formando (com o artigo) um sintagma nominal? Ou deveremos separá-los em classes diferentes, já que só *branco*, e não *casa*, concorda em gênero com o núcleo do sintagma a que pertence?

Esse tipo de problema é tão generalizado, como sabem os que já se ocuparam criticamente de gramática, que talvez valha a pena rever fundamentalmente nossa idéia de classe de palavras. Vou desenvolver essa questão adiante, propondo uma solução em linhas comuns em Lingüística, ainda que desconhecidas da GT.

Peguemos como exemplos as três palavras seguintes:

casa, ele e branco. Examinaremos seu comportamento de acordo com três critérios sintáticos, suficientes, como se verá, para distingui-las umas das outras. O primeiro crité rio (ou traço) é o dado acima, em (56): a propriedade de poder ser núcleo de um sintagma nominal. Já sabemos que as três palavras em questão são marcadas positivamente quanto a esse traço: ou seja, todas as três *podem* ser núcleos de sintagmas nominais (ver os exemplos de (57) a (59)). Assim, poderemos começar a representar a classificação dessas palavras da seguinte forma:

(61) traço casa ele branco

(56) (= "poder ser núcleo de um sintagma nominal")

Por enquanto, pois, as três palavras examinadas se comportam de maneira idêntica (em oposição a *sem* ou *corremos*, que devem ser marcadas "—" para esse traço). Parece, por conseguinte, que deveremos colocá-las na mesma classe. Mas as coisas mudam, assim que acrescen tamos um segundo traço, digamos:

(62) A propriedade de ocorrer após um artigo, for mando a seqüência um sintagma nominal.

Verificaremos facilmente que *casa* e *branco* devem ser marcados positivamente quanto a (62), mas ele precisa ser marcado negativamente, pois não pode ocorrer após artigo.

À primeira vista não parece haver problema. Pode ríamos descrever esses fatos de maneira mais ou menos tradicional, assim: Há uma classe de palavras definida pela sua propriedade de poder ser núcleo de sintagmas nominais; essa classe inclui *casa, ele e branco*. Essa classe se subdivide em duas subclasses, segundo o critério de po der ou não ocorrer após artigo. *Casa* e *branco* ficam em uma subclasse, *ele* fica na outra.

Mas os fatos não são simples assim. Ao considerarmos outros traços, veremos que as classes, longe de se deixarem dividir e subdividir nitidamente e sem sobreposições, ten dem a misturar-se, de tal modo que as subclassificações se cruzam. Para ilustrar isso, passemos ao terceiro traço:

(64) A propriedade de concordar em gênero com o núcleo do sintagma a que pertence .

Casa e *ele* serão marcados negativamente (pois não concordam); *branco* será marcado positivamente.

Uma solução que ocorre é ver aí três classes, dando-se um nome a cada uma, ainda mais que tais nomes estão disponíveis na nomenclatura tradicional: “substantivo”, “pronome pessoal”, “adjetivo”. Isso equivale a definir cada uma dessas classes de maneira complexa, através de um conjunto de vários traços, e não de um só traço. A conveniência dessa solução depende de os “cruzamentos” de traços não serem muito extensos na língua. Explico-me melhor; a questão a ser formulada é a seguinte: Pode-se dizer que muitos traços, ou os traços mais importantes, coincidem quanto ao corte que impõem ao conjunto dos itens léxicos? Em que medida diferentes traços, ao serem utilizados como base para a divisão do léxico em grupos, nos fornecem agrupamentos semelhantes de palavras? Se acontecer que haja muita coincidência entre as classificações impostas por muitos traços diferentes, então fará sentido falar em “classes” mais ou menos à maneira tradicional. Nesse caso, termos como “substantivo” ou “verbo” serão usados como abreviaturas para matrizes típicas de traços. E, conversamente, a definição de cada uma dessas classes será uma matriz de traços do tipo do exemplificado em (65), embora, certamente, mais complexa.

Se, ao contrário, encontrarmos que não há grandes linhas de coincidência entre os cortes definidos na massa dos itens léxicos pelos diferentes traços, então termos como “substantivo” etc, perderão significatividade para a análise. Por isso, é importante averiguar se os fatos da língua reve lam ou não tais coincidências entre os traços. Essa é uma das muitas direções de pesquisa suscitadas por este novo ponto de vista quanto à classificação das palavras.

Além do problema da coincidência entre os traços quanto aos cortes impostos ao conjunto das palavras, men cionarei outro, também relevante. Referi-me acima aos tra ços “mais importantes”, cuja coincidência seria mais

signi ficativa para o estabelecimento das classes. Mas que crité rios temos para avaliar a importância relativa de um traço? Esse é outro problema que terá de ser atacado: o da hierarquização dos traços segundo sua importância para a análise.

A importância da hierarquização dos traços vem de que dela depende o estabelecimento adequado das classes de palavras. Assim, é possível classificar as palavras de tal modo que as classes não tenham quase utilidade ne nhuma — justamente porque os traços escolhidos como critérios de classificação são gramaticalmente pouco signi ficativos. Por exemplo, digamos que vamos fazer o primeiro grande corte no léxico do português, iniciando a divisão das palavras em classes; e digamos que escolhemos como critério o traço (56), "poder ser núcleo de um sintagma nominal". Observaremos, em primeiro lugar, que esse traço divide o léxico português em dois grandes grupos, de tamanhos comparáveis. De um lado, marcadas positivamente, estarão palavras como *casa, branco, ele,* e muitas outras. De outro, marcadas negativamente, *estarão corremos, pois, de, rapidamente* etc. O que faz dessa divisão um fato importante da língua é que o traço (56) tem cor relação com outros traços. Por exemplo, as palavras mar cadas positivamente quanto a esse traço não somente podem ser núcleos de sintagmas nominais, mas também são as únicas palavras que podem reger algum tipo de concordância: encontramos na língua casos em que um termo concorda com *casa*, ou com *branco*, ou com *ele*; mas nunca encontraremos um termo concordando com uma palavra do segundo grupo (*corremos, pois, de, rapidamente* etc.).

Comparemos agora o caso do traço (56) com o de um outro traço, assim definido:

(66) A propriedade de ocorrer obrigatoriamente precedido de artigo definido, quando dentro de um sintagma nominal ".

Esse traço divide os itens léxicos da língua em dois grupos muito desiguais em tamanho. De um lado, marca das negativamente, ou seja, *não* exigindo artigo, fica a imensa maioria das palavras da língua. Do outro lado, marcados positivamente, fica uma dúzia de nomes pró prios como *o Cairo, o Rio, o Leblon, a Barroca, a Penha*, todos nomes de lugares. É fácil ver como o traço (66) tem pouca importância no conjunto da gramática, não só por interessar a poucos itens, mas principalmente porque não há outras afirmações gramaticais que se refiram às classes que ele define: palavras como *Cairo* e *Barroca*, à parte o fato de exigirem o artigo definido, não têm nada de característico, gramaticalmente falando.

Dessas observações, podemos concluir com segurança que vale a pena considerar o traço (56) mais importante do que o traço (66); conseqüentemente, (56) deverá de sempenhar um papel de primazia na definição das classes de palavras, enquanto (66) poderá ser ignorado ou, quan do muito, relegado à descrição de simples detalhes.

Uma vez estabelecidas com certa segurança as principais classes de palavras, a grande maioria dos itens deve encaixar-se sem problemas maiores em uma ou outra classe.

Mas certamente restará um resíduo de palavras que não se colocam com facilidade em nenhuma das classes estabelecidas. Veremos que esse problema, insolúvel dentro do sistema tradicional de classificação, pode ser tratado sem dificuldade no sistema de traços, dada a maior flexibilidade (e complexidade) deste último. Neste caso, aliás, a complexidade não é um mal, pois os fatos parecem ser complexos: complexos demais para a simplicidade da abordagem tradicional.

Consideremos o caso de *alto* em oposição a *bom*. De certo ponto de vista, essas duas palavras são muito semelhantes, isto é, compartilham muitos traços. Por exemplo, ambas são marcadas positivamente para (56) (ou seja, podem ser núcleos de sintagmas nominais); positiva mente para (62) (podem ocorrer após artigo); positiva mente para (64) (concordam em gênero); negativamente para a concordância em pessoa etc. Por isso, acabaram sendo colocadas na mesma classe pela GT: são “adjetivos”. É vantajoso para a análise conservar o essencial dessa intuição, certamente correta. Mas isso não significa que as duas palavras sejam sintaticamente idênticas. A palavra *alto* pode ocorrer em certos ambientes dos quais *bom* está excluído, como em

(67) ela fala muito alto

(68) ela fala muito bom (Este exemplo não é gramatical).

Esse ambiente não é exclusivo de *alto*; existe todo um grupo de palavras que ocorrem nas mesmas circuns tâncias, e que não obstante se diferenciam de *alto* por outros traços: palavras como *seriamente, cuidadosamente, bem* etc. O traço que essas palavras têm em comum com *alto* pode ser formulado assim:

(69) A propriedade de ocorrer em um sintagma com um verbo não de ligação, sem concordar com nenhum termo 12.

Alto, portanto, se classifica juntamente com *bom*, segundo a maioria dos critérios; mas, segundo um critério, se classifica junto com *seriamente*. Em termos da GT, essa é uma situação insolúvel, O máximo que se faz é introduzir afirmações do tipo “o adjetivo x funciona aqui como advérbio”, afirmações que, se levadas a sério, acabarão destruindo a própria noção de classe de palavras. Sabemos que (pelo menos segundo a DGImp.) as classes se definem em função de seu comportamento gramatical; assim, um adjetivo é uma palavra que se comporta (ou “funciona”) como adjetivo — como quer que isso se defina. Nesse contexto, dizer de um adjetivo que funciona como advér bio é simplesmente uma contradição em termos.

A presente proposta, por outro lado, permite enumerar independentemente os traços de cada uma das palavras, como mostra a matriz abaixo:

Não há dúvida de que a introdução de matrizes como no (70) para caracterizar plenamente as propriedades sintáticas dos itens léxicos representa uma complicação face ao sistema tradicional de classes. Entretanto, como fiz notar, isso não é objeção válida ao sistema de traços, visto que o sistema tradicional não dá conta dos fatos da língua. O sistema de classificação por traços permite o tratamento de casos que, em um

sistema de classes mutuamente exclusivas, não poderiam ser analisados convenientemente. A maior parte dos problemas estão por resolver; mas acre dito que as colocações aqui feitas contribuirão para dar uma direção à pesquisa. As grandes tarefas iniciais se resumem em (a) definir os traços classificatórios; (b) estabelecer entre eles uma hierarquia de importância gramatical; e (c) apurar o grau de coincidência dos cortes impostos ao conjunto dos itens léxicos pelos traços mais importantes (e, daí, apurar até que ponto valerá a pena continuar falando de "classes de palavras", já agora definidas como uma maneira abreviada de se referir a certas matrizes típicas de traços). A partir daí é que poderemos criar uma classificação realmente funcional das palavras do português, fazendo justiça a toda a sua complicação e riqueza.

5 Os dados da análise

Vamos ficar por aqui. A tarefa de lançar as bases da nova gramática é necessariamente longa, e ultrapassa largamente os limites do presente texto. Espero ter abordado alguns dos temas mais importantes; e se a leitura deste livro vier a incentivar pesquisas no sentido de criar uma nova linguagem para a gramática pedagógica, ele terá cumprido sua finalidade.

Antes de encerrar, vou considerar brevemente um último ponto, não diretamente ligado à doutrina gramatical propriamente dita, mas relevante por ser um pré-requisito a qualquer descrição do português padrão: o problema da seleção dos dados a serem descritos.

Trata-se, naturalmente, de descrever o português-padrão do Brasil; mas onde encontramos textos vazados nessa variedade da língua? Examinemos uma revista noticiosa semanal; um grande jornal diário; um compêndio de Física; o texto de uma lei qualquer; um romance de Guimarães Rosa; poemas de Mário de Andrade; e a partes dialogadas de um romance de Jorge Amado. Todos são exemplos do português escrito e são, por definição, amostras de português-padrão. Mas certamente há muitas diferenças gramaticais (e léxicas) entre eles. Por exemplo, em Jorge Amado encontraremos grande número de formas gramaticais típicas da variedade coloquial do português; em Mário de Andrade também, mas talvez com algumas diferenças de ordem regional; o texto jurídico, por outro lado, apresentará possivelmente construções de sabor arcaico; Guimarães Rosa inclui formas e construções *sui generis*, proposital mente criadas para obter efeitos estéticos.

Se quisermos considerar todos esses textos como exemplos da variedade-padrão da língua, poderemos ver-nos frente a dilemas difíceis de solucionar: a construção *vou lá não* deve ser acolhida como padrão? (Jorge Amado a emprega com freqüência.) Que dizer do tratamento você com o possessivo vosso? (É encontrado em Mário de Andrade.) E se decidirmos excluir essas construções, como o justificaremos? Se argumentarmos que *vou lá não* é coloquial, em que definição de "coloquial" nos apoiamos? Pois, afinal, "coloquial" é, por definição, a forma que ocorre na linguagem falada espontânea, e *vou lá não* aparece no texto de um romance. O exemplo de Mário de Andrade é ainda mais difícil de excluir. Deveríamos talvez lançar

mão de um critério estatístico, deixando de fora o uso de você com vosso por sua raridade?

Acredito que a solução está em outra parte. Deixemos de lado, por ora, os textos literários e voltemo-nos para o outro tipo de textos: a linguagem técnica e jornalística, encontrada, por exemplo, nas revistas semanais, nos jornais, nos livros didáticos e científicos. Ora, examinando esses textos, encontraremos uma grande uniformidade gramatical: não só as formas e construções encontradas nos jornais e revistas são as mesmas dos compêndios e livros científicos, mas também não se percebem variações regionais marcadas: um jornal de Recife usa sensivelmente a mesma língua de um jornal de Porto Alegre ou de Cuiabá. Isto é, existe um português-padrão altamente uniforme do País; e podemos contar encontrá-lo nos textos jornalísticos e técnicos 13.

É claro que também encontramos obras literárias escritas estritamente segundo esse padrão; mas não podemos contar com isso *a priori*: muitas obras literárias fogem a ele, utilizando variedades coloquiais ou mesmo pessoais. Sou de opinião de que os dados que fundamentarão a gramática devem ser retirados desse padrão técnico-—jornalístico.

A primeira razão é que tais textos oferecem uma uniformidade de estrutura que nos permite elaborar a descrição com maior coerência. Torna-se possível excluir construções como *vou lá não*, ou então o *meu/teu corpo-fruto* (Ronald Cláver), pois não ocorrem no corpus que selecionamos.

A segunda razão é que a própria linguagem literária pode ser (e na prática costuma ser) estudada em termos de desvio a partir de um padrão básico. O autor utiliza construções coloquiais para criar um efeito específico, ou viola intencionalmente as regras gramaticais também com objetivo estético. Ora, tudo isso pode ser estudado em relação com o padrão básico; logo, se selecionamos esse padrão como a fonte de nossos dados, estaremos ao mesmo tempo lançando uma base para o estudo gramatical dos textos literários.

Finalmente, há uma terceira razão, talvez a mais importante. Argumenta-se que uma das finalidades do en sino gramatical é conscientizar o estudante de sua língua, da língua que ele deve aprender a manejar, seja lendo, seja escrevendo. Mas certamente muito poucos estudantes

13 Não se veja aqui contradição com o que foi dito no Prefácio, que o português do Brasil não é “uma entidade simples e homogênea”. A heterogeneidade se manifesta amplamente nas variedades coloquiais (regionais e sociais), assim como na linguagem literária. Auniformidade de que falo se refere a textos escritos não-literários (com exceções especiais, como a de alguns textos jurídicos).

Chegarão a produzir textos literários; digo mais: poucos chegarão a adquirir o hábito de ler textos literários. Mas é certamente necessário (embora ainda estejamos terrívelmente longe de consegui-lo) que eles cheguem a manejar a linguagem técnica e jornalística, pelo menos como leitores. Uma exigência que a civilização faz do homem moderno é que seja capaz de adquirir conhecimentos técnicos a partir de textos escritos (mesmo

que seja apenas como montar uma estante); e que seja capaz de obter informações cruciais para sua participação na vida política de sua comunidade e do País — informações que lhe chegam, principalmente, através de material escrito: jornais, revistas, comunicados, manifestos.

Não pretendo com essas considerações diminuir a importância do estudo da literatura, ou do domínio da linguagem literária. Mas acho que, em um sistema educacional que tão freqüentemente deixa de alcançar o mínimo, é importante colocar objetivos mais realistas para o ensino da língua.

Por razões como essas que gostaria de sugerir que a gramática seja (pelo menos em um primeiro momento) uma descrição do português-padrão tal como se manifesta na literatura técnica e jornalística.

6 Vocabulário crítico

Apesar da preocupação que tive de utilizar uma terminologia tanto quanto possível conhecida do eventual leitor destas páginas, não me foi possível evitar de todo o uso de termos técnicos menos familiares. Alguns destes se en contram definidos no próprio texto; quanto aos restantes, listei-os abaixo, com as respectivas definições (que se deve entender como muito sumárias, dirigidas apenas às neces sidades imediatas da compreensão deste livro). Acrescentei a lista das abreviaturas introduzidas no texto, para facilidade de referência.

Dêixis (adj. dêitico): chamam-se dêiticos os elementos lingüísticos cuja referência só se pode determinar com base no contexto extralingüístico. Por exemplo, a pala vra eu, ou a palavra isso em: isso é muito caro para mim.

DGEx.: Doutrina Gramatical Explícita.

DGImp.: Doutrina Gramatical Implícita.

Gramática: conjunto de regras que definem as combinações possíveis dos elementos léxicos de uma língua, assim como sua interpretação semântica e sua pronúncia.

GT: Gramática Tradicional.

Léxico: lista de todas as palavras (ou de todos os mor femas) de uma língua, juntamente com as características próprias, idiossincráticas, de cada uma.

Nódulo vazio: elemento abstrato correspondente a um constituinte (por exemplo, a uma palavra) e que não tem realização fonológica.

Referência: relação que existe entre uma expressão lingüística e o objeto (concreto ou abstrato) que essa ex pressão nomeia. Assim, diz-se que a referência da expressão lingüística Carlinhos é a relação entre essa palavra e a pessoa chamada "Carlinhos"; e a referência de meu pé na frase meu pé está inchado é a relação entre esse sintagma e o pé de quem enunciou a frase. Nem todas as expressões lingüísticas são referenciais; algumas apenas descrevem um possível objeto, inclusive sem compromisso quanto à sua existência. Assim, em encontrei um charuto de 16 cm, o sintagma um

charuto de 16 cm é referencial: nomeia um objeto determinado, tomado como existente. Já em eu gostaria de encontrar um charuto del6 cm, o mesmo sintagma é não-referen cial; o falante est apenas descrevendo os atributos de um objeto sem o tomar como existente.

Sintagma nominal: qualquer seqüência de elementos da língua que possa desempenhar certas funções sintáticas, como a de sujeito, a de objeto direto, a de regime de uma preposição etc.

SVV: Solicitação de Valor de Verdade.

Traço (sintático): propriedade sintática de uma palavra ou de um sintagma. Por exemplo, a propriedade de poder ser o núcleo de um sujeito; ou a de pertencer a um paradigma, cujos membros se opõem quanto à pessoa, tempo e número.

Bib1iografia comentada

ARONOFF, Mark (1976). *Word formation in generative grammar.* Cambridge, Mass., MIT Press. Princípios de morfologia, segundo a teoria gerativa.

BECHARA, Evanildo (1968). *Moderna gramática portuguesa.* S. Paulo, Nacional. Texto bem realizado, mas dentro dos limites da gramática tradicional brasileira.

CAMARA, J. Mattoso (1964). *Princípios de Lingüística geral.* Rio de Janeiro, Acadêmica. A velha obra de Mattoso Camara (a 1.a edição é de 1942) ainda merece ser consultada, especialmente os capítulos de fonética, fonologia e morfologia.

CHAFE, Wallace L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In: Li, Charles N. ed., *Sub ject and topic.* N. York, Academic Press. Artigo básico sobre o problema dos *status* funcionais.

* As obras assinaladas são de leitura difícil, exigindo certo conhecimento prévio de teorias lingüísticas.

CUNHA, Celso (1975). *Gramática da língua portuguesa.* Rio de Janeiro, MEC-Fename. A mais conhecida, e talvez a melhor, das gramáticas portuguesas atuais.

FULGÊNCIO, Lúcia (1983). *O problema da interpretação dos elementos anafóricos.* Dissertação de Mestrado, UFMG. Discussão das condições de uso dos elementos anafóricos. Há um resumo em *Ensaios de Lingüística 9* (revista do Departamento de Lingüística e Teoria Literária da UFMG, 1983).

GARCIA, Othon M. (1975). *Comunicação em prosa moderna.* Rio de Janeiro, Fundação Getúlio Vargas. Trabalho interessante, não só por demonstrar independência frente a certos dogmas tradicionais, mas também por estudar assuntos geralmente negligenciados, como o discurso.

GRUBER, Jeffrey S. *(1965). Studies in lexical relations.* Dissertação de Doutorado, MIT. Trabalho pioneiro no campo das relações temáticas (“agente”, “paciente” etc.).

HAUY, Amini B. (1983). *Da necessidade de uma gramática-padrão da língua portuguesa. S.* Paulo, Ática. Crítica lúcida e bastante severa da lógica subjacente à gramática tradicional.

* JACKENDOFF, Ray S. (1972). *Semantic interpretation in generative grammar.* Cambridge, Mass., MIT Press. Obra básica de semântica da sentença, seguindo a perspectiva gerativista.

JESPERSEN, Otto (1924). *The philosophy of grammar.* London, Allen & Unwin. Livro antigo e um pouco difuso, mas cheio de intuições interessantes sobre a gramática.

* LEMLE, Minam (1982). *Análise sintática. S.* Paulo, Ática. Trabalho importante de aplicação da teoria gerativa *à* análise do português.

LIBERATO, Yara O. (1980). *Sobre a oposição dado/novo.* Dissertação de Mestrado, UFMG. Apanhado útil do problema dos *status funcionais.* Amplia e em parte revisa Chafe (1976).

LYONS, John (1977). *Semantics,* Cambridge, England, Cambridge University Press. Boa introdução geral à semântica. Há tradução portuguesa (São Paulo, Martins Fontes, 1980).

* MARTINET, André (1979). *Grammaire fonctionnelle du français.* Crédif — Didier. Tentativa de análise do francês segundo uma teoria estruturalista (a variedade “funcionalista” defendida pelo Autor).

* NIDA, Eugene (1949). *Morphology;* the descriptive analysis of words. Ann Arbor, University of Michigan.

Obra típica da morfologia estruturalista.

* PONTES, Eunice (1973). *Estrutura do verbo no português coloquial.* Petrópolis, Vozes. Importante análise da morfologia verbal portuguesa. Também de interesse é o capítulo sobre fonologia.

QUIRK, Randolph; GREENBAUM, Sidney; LEECH, Geoffrey; SVARTVIK, Jan (1972). *A grammar of contemporary English.* London, Longman. Tentativa de atualização da gramática inglesa. Contém informações de interesse, e é de leitura mais fácil do que o trabalho de Martinet.

SAPIR, Edward (1921). *Language;* an introduction to the study of speech. N. York, Harcourt Brace Jovanovich. Um clássico. Leitura recomendada para qualquer interessado em linguagem. Há tradução portuguesa (Rio de Janeiro, Acadêmica, 1971).